A Multiscale Forecasting Methodology for Power Plant Fleet Management

A Thesis
Presented to
The Academic Faculty

by

Hongmei Chen

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Aerospace Engineering
Georgia Institute of Technology
January 2005

A Multiscale Forecasting Methodology for Power Plant Fleet Management

Approved by:

Dr. Dimitri Mavris, Advisor
Aerospace Engineering
Georgia Institute of Technology

Dr. Daniel Schrage
Aerospace Engineering
Georgia Institute of Technology

Dr. Brani Vidakovic
Industrial System Engineering
Georgia Institute of Technology

Mr. Mark Waters
Aerospace Engineering
Georgia Institute of Technology

Dr. Vitali Volovoi
Aerospace Engineering
Georgia Institute of Technology

Mr. Mike Sullivan
Senior Application Engineer
GE Power Systems

Date Approved: February 2005

To my husband Yanwu Yin

To my parents

ACKNOWLEDGEMENTS

A Journey Is Easier When You Travel Together.

This dissertation is the result of almost four years of work during which I have been accompanied and supported by many people. It is with great pleasure that I now have the opportunity to express my gratitude to all of them.

The first person to whom I would like to express my sincere thanks and appreciation is my advisor, Dr. Dimitri Mavris. His overt enthusiasm and integral view on research have made a deep impression on me. I owe him enormous gratitude for his believing in my potential, for giving me various chances, and for guiding me in my research. He cannot imagine the extent to which I have learned from him. I am thankful that I had the opportunity to know and learn from Dr. Mavris.

I would also like to express my deep gratitude to Dr. Brani Vidakovic for guiding me in the world of time series and nonparametric statistics, for his willingness to share his ideas in research problems, and for the energy he put into advising me on my dissertation work. Besides being an excellent supervisor, Dr. Vidakovic has been a close friend.

Special thanks go to Dr. Schrage, Mr. Waters, Dr. Volovoi, and Mr. Sullivan for their valuable comments and ideas, continued encouragement, and support throughout my research. I also thank them for taking the time to read and provide feedback on this dissertation.

I am also grateful to various researchers for taking the time to guide me through my dissertation. I would also like to thank those who gave me their valuable time, skills, and enthusiasm during these years, particularly the members of Aerospace System Design Laboratory (ASDL).

Finally, I would like to thank my parents, Mr. Guangcheng Chen and Mrs. Shuqin Wang, for their continued support throughout my education at Georgia Tech. I would also like to acknowledge the warm support and caring of my dear husband, Yanwu Yin.

TABLE OF CONTENTS

DEDICATION . . . . **iii**

ACKNOWLEDGEMENTS . . . . **iv**

LIST OF TABLES . . . . **ix**

LIST OF FIGURES . . . . **xi**

LIST OF ABBREVIATIONS . . . . **xvii**

SUMMARY . . . . **xix**

I MOTIVATION **1**

1.1 Multi-Timescale Decision Making 4

1.2 The Forecasting Problem for Power Plants 7

1.3 External Information Adaptive Processing 12

1.4 Research Questions and Assumptions 16

1.4.1 Research Questions 16

1.4.2 Hypotheses 17

II BACKGROUND **18**

2.1 Major Power Plant Decision Actions 18

2.1.1 Optimal Dispatch 18

2.1.2 Unit Commitment 20

2.1.3 System Maintenance Scheduling 23

2.1.4 System Operational Planning 27

2.1.5 System Capacity Expansion 30

2.2 Unit and System Maintenance Constraints 33

2.2.1 Unit Maintenance Constraints 34

2.2.2 System Maintenance Constraints 35

2.3 Components Fired Factored Hours and Fired Factored Starts 36

2.3.1 Inspections 36

2.3.2 Duties 37

2.3.3 Fired Factor Hours/Starts 39

2.4 Forecasting Variables 43

2.4.1 Electric Market . . . 43
2.4.2 Customer Demand Forecasting . . . 44
2.4.3 Electricity Spot-Market Price Forecasting . . . 46
2.4.4 Fuel Requirement Forecasting . . . 49
2.5 Current Forecasting Methods . . . 51
2.5.1 Qualitative Forecasting Methods . . . 52
2.5.2 Time Series Forecasting Methods . . . 54
2.5.3 Casual Forecasting Methods . . . 56
2.5.4 Simulation Methods . . . 57
III APPROACH . . . 59
3.1 Power Plant Fleet Management . . . 60
3.1.1 Modeling and Simulation Environment . . . 60
3.1.2 Unit Operating Conditions . . . 61
3.1.3 System Characteristics . . . 63
3.1.4 Identify Time Scales . . . 65
3.1.5 Determine the System Operating Strategy . . . 68
3.1.6 Determine the System Maintenance Schedule . . . 71
3.1.7 Investigate the System Capacity Expansion Plan . . . 74
3.2 Analysis of Electric Market Dynamics . . . 76
3.2.1 Fourier Transform . . . 77
3.2.2 Multi-Resolution Analysis . . . 79
3.2.3 Wavelet Transform . . . 81
3.2.4 Wavelet Families . . . 91
3.3 Forecasting Method - WAW . . . 93
3.3.1 Forecasting Methodology . . . 93
3.3.2 Forecasting Error Analysis . . . 97
3.3.3 Block Bootstrapping Estimate of the LCC . . . 105
3.4 Uncertainty Exploration . . . 111
3.4.1 External Factors Identification . . . 115
3.4.2 Scenarios Generation . . . 116
3.4.3 Scenarios Analysis . . . 118

IV FORECASTING RESULTS AND ANALYSIS 120

4.1 Customer Demand Forecasting 120

4.1.1 Historical Data 120

4.1.2 Data Analysis 122

4.1.3 Forecasting Results 128

4.2 Natural Gas Prices Forecasting 129

4.2.1 Historical Data 129

4.2.2 Data Analysis 130

4.2.3 Forecasting Results 138

4.3 Electricity Prices Forecasting 139

4.3.1 Historical Data 139

4.3.2 Data Analysis 140

4.3.3 Forecasting Results 146

4.4 Forecasting Errors 147

4.5 Comparisons With Holt-Winters' Method 149

V POWER PLANT FLEET MANAGEMENT 152

5.1 Unit Conditions and System Characteristics 152

5.1.1 Unit Load Settings 152

5.1.2 System Capacity 152

5.1.3 Economical Operating Period 153

5.1.4 Operation Profile 154

5.1.5 Operating Condition Ranking 154

5.2 System Operating Strategies And System Maintenance Schedules 155

5.2.1 Baseline SMS and SOS 156

5.2.2 Deviation Analysis 161

5.3 System Capacity Expansion Plans 179

5.4 A Bootstrapping Estimate of the LCC 182

5.5 Uncertainty Exploration 183

VI CONCLUSIONS 212

6.1 Conclusions 212

6.2 Future Work and Recommendations 218

APPENDIX A — THE COMPUTATIONS OF MAINTENANCE FACTORS **222**

REFERENCES **226**

VITA **234**

LIST OF TABLES

1 FA Gas Turbine Typical Operational Duties 38
2 Energy at Each Level and the Recovered data 98
3 Tests for White Noise 100
4 Energy at Each Level and the Recovered Data 100
5 Energy at Each Level and the Recovered data 103
6 Energy at Each Level and the Recovered data 105
7 Second Level Harmonic Regression Coefficients 125
8 Third Level Harmonic Regression Coefficients 127
9 Second Level Harmonic Regression Coefficients 134
10 Second Level Upper Envelop Gaussian Regression Coefficients 134
11 Second Level Bottom Envelop Gaussian Regression Coefficients 135
12 Third Level Harmonic Regression Coefficients 136
13 Third Level Upper Envelop Gaussian Regression Coefficients 136
14 Third Level Lower Envelop Gaussian Regression Coefficients 138
15 Second Level Harmonic Regression Coefficients 144
16 Third Level Harmonic Regression Coefficients 146
17 Forecasting Errors 149
18 Normalized Generation Unit Output 152
19 System Capacity and Available Capacity 153
20 Continuous Operation Profile 154
21 Operating Condition Ranking 155
22 Operating Condition vs. Color 156
23 Baseline: Maintenance Activities in the 4^{th} Quarter 158
24 Baseline: System Status Adjustments in the 4^{th} Quarter 159
25 Baseline: Maintenance Activities in the 14^{th} Quarter 159
26 Baseline: System Status Adjustments in the 14^{th} Quarter 160
27 Deviation 1: Unscheduled Maintenance 161
28 Deviation 1: Maintenance Activities in the 4^{th} Quarter 162
29 Deviation 1: System Status Adjustments in the 4^{th} Quarter 164

30 Deviation 1: Maintenance Activities in the 14^{th} Quarter 164
31 Deviation 1: System Status Adjustments in the 14^{th} Quarter 165
32 Deviation 2: Unscheduled Maintenance 165
33 Deviation 2: Maintenance Activities in the 14^{th} Quarter 166
34 Deviation 2: System Status Adjustments in 14^{th} Quarter 167
35 Deviation 3: Unscheduled Maintenance 168
36 Deviation 3: Maintenance Activities in the 4^{th} Quarter 169
37 Deviation 3: System Status Adjustments in the 14^{th} Quarter 169
38 Deviation 3: Maintenance Activities in the 14^{th} Quarter 170
39 Deviation 3: System Status Adjustments in the 14^{th} Quarter 171
40 Deviation 4: Unscheduled Maintenance 171
41 Deviation 4: Maintenance Activities in the 4^{th} Quarter 172
42 Deviation 4: System Status Adjustments in the 4^{th} Quarter 173
43 Deviation 5: Unscheduled Maintenance 174
44 Deviation 5: Maintenance Activities in the 14^{th} Quarter 175
45 Deviation 5: System Status Adjustments in the 14^{th} Quarter 176
46 Deviation 6: Unscheduled Maintenance 177
47 Deviation 6: Maintenance Activities in the 14^{th} Quarter 177
48 Deviation 6: System Status Adjustments in the 14^{th} Quarter 178
49 Expansion: Normalized Generation Unit Output 180
50 LCC for Each Pseudo Sample 183
51 Morphological Fields For Parameters 184
52 Total Cost for Each Scenario 193
53 Total Cost for Each Scenario 203

LIST OF FIGURES

1 Time Scales of Major Decision Actions 5

2 Forecasting Process 8

3 Time Value of Information 15

4 Optimal Dispatch Decision Action 20

5 Unit Commitment Decision Action 23

6 System Maintenance Scheduling Decision Action 26

7 Load Profiles for Four Seasons 27

8 Contributing Factors to the Volatility of Customer Demands 28

9 System Operation Planning Decision Action 30

10 System Expansion Planning Decision Action 33

11 Trend of Maintenance Cost 34

12 Different Inspections Work Scope 37

13 GE Bases Gas Turbine Maintenance Requirements on Independent Counts of Starts and Hours 39

14 Estimated Repair and Replacement Cycles 40

15 GE Maintenance Interval for Hot-Gas-Path Inspections 41

16 Bucket Life Firing Temperature Effect 42

17 Historical Monthly Customer Demand 46

18 Historical Monthly Electricity Prices 48

19 Factors Contributing to Cost of Electricity 49

20 Fraction of Fuel Cost in the Total LCC of a Power Plant 50

21 Historical Monthly Natural Gas Prices 51

22 Flow Chart of the Modeling Methodology 60

23 Load Setting and Firing Temperature Relationship for Simple Cycle Operation and Heat Recovery Operation 62

24 Operating Conditions 63

25 Frequency-Time Domain of Wavelet Transform 82

26 Time-Frequency Tiles and Coverage of the Time-Frequency Plane 83

27 Decomposition Algorithm 87

28 Reconstruction Algorithm 89

29 The Discrete Wavelet Transform Lacks Translation-Invariance 90

30 Data Series: Doppler 92

31 Doppler in the Wavelet Domain through the DWT 92

32 Wavelet Transform by the DWT and the NDWT 93

33 Wavelet Families (a) Haar (b) Daubechies4 (c) Coiflet1 (d) Symmlet2 (e) Meyer (f) Morlet (g) Mexican Hat 93

34 Decimated and Non-decimated Wavelet Transforms 94

35 QQ Plot of Sample Data versus Standard Normal 99

36 QQ Plot of Sample Data versus Standard Normal 100

37 AR Model using Different Wavelet Filters 102

38 AR Model for Time Series of Different Lengths 104

39 AR Model for Time Series of Different Lengths with Randomly Generated Variance 106

40 Data and Histogram 108

41 Block Bootstrap Process 111

42 Morphological Field 116

43 Morphological Fields for Parameters 117

44 Residential and Commercial Demand ($Tbtu$) 121

45 Seasonal Patterns Existing in the Historical Data 121

46 The Log Transform of Residential and Commercial Demand 122

47 Customer Demand in the Wavelet Domain, Performed with Symmlet (8) . . 123

48 The First Level Data and Fitness Test 124

49 The First Level Predicted by the AR(8) Process 124

50 The Second Level Fitted Using Harmonic Regression ($\omega = 0.5244$) 125

51 Second Level Forecasting Results ($\omega = 0.5244$) 126

52 The Third Level Fitted Using Harmonic Regression ($\omega = 0.5174$) 126

53 Third Level Forecasting Results ($\omega = 0.5174$) 127

54 Fourth Level Forecasting Results Performed by Holt-Winters' Method . . . 128

55 Forecasting Results for the Following 24 Months 128

56 Natural Gas Electric Utility Purchase Prices (cnt/mcf) 129

57 Natural Gas Prices in the Wavelet Domain, Performed with Symmlet (8) . 130

58 The First Level Data and Fitness Test 131
59 The First Level Data and the External Factor IDPlot 131
60 Correlation between the First Level Data and the External Factor 132
61 Correlation Relationship of the Residuals 133
62 The First Level Data Fitted Using ARMAX Process 133
63 The Second Level Data Fitted Using Harmonic Regression ($\omega = 0.5233$) . . 134
64 Second Level Upper Envelop Fitted Using Gaussian Regression 135
65 Second Level Bottom Envelop Fitted Using Gaussian Regression 135
66 Second Level Forecasting Results ($\omega = 0.5233$) 136
67 The Third Level Fitted Using Harmonic Regression ($\omega = 0.520$) 137
68 Third Level Upper Envelop Fitted Using Gaussian Regression 137
69 Third Level Bottom Envelop Fitted Using Gaussian Regression 137
70 Third Level Forecasting Results ($\omega = 0.520$) 138
71 Fourth Level Forecasting Results by Holt-Winters' Method 139
72 Forecasting Results for the Following 24 Months 139
73 Electricity Industrial Sector Prices ($hcnt/kwh$) 140
74 Seasonal Patterns Existing in the Historical Data 140
75 Electricity Prices in the Wavelet Domain, Performed with Symmlet (8) . . . 141
76 The First Level Data and Fitness Test 141
77 The First Level Data and the External Factor IDPlot 142
78 Correlation between the First Level Data and the External Factor 142
79 Correlation Relationship of the Residuals 143
80 The First Level Data Fitted Using ARMAX Process 143
81 The Second Level Fitted Using Harmonic Regression ($\omega = 0.5254$) 144
82 Second Level Forecasting Results ($\omega = 0.5254$) 145
83 The Third Level Fitted Using Harmonic Regression ($\omega = 0.5211$) 145
84 Third Level Forecasting Results ($\omega = 0.5211$) 146
85 Fourth Level Forecasting Results by Holt-Winters' Method 147
86 Forecasting Results for the Following 24 Months 147
87 Customer Demand Validation ($Tbtu$) 148
88 Electricity Price Validation ($hcnt/kwh$) 148

89 Natural Gas Price Validation (cnt/mcf) 149
90 Residential and Commercial Demand ($Tbtu$) 150
91 Electricity Price Comparison ($hcnt/kwh$) 150
92 Natural Gas Price Comparison (cnt/mcf) 151
93 Economical Operating Period 153
94 System Status vs. Color 155
95 Baseline: SOS and SMS 157
96 Baseline: System Generation vs. Customer Demand 158
97 Baseline: System Reactions in the 4^{th} Quarter 159
98 Baseline: System Reactions in the 14^{th} Quarter 160
99 Baseline: Power Plant Cost Distributions 161
100 Deviation Locations in the Baseline Operation 162
101 Deviation 1: SOS and SMS 163
102 Deviation 1: System Reactions in the 4^{th} Quarter 163
103 Deviation 1: System Reactions in the 14^{th} Quarter 164
104 Deviation 1: Power Plant Cost Distributions 165
105 Deviation 2: SOS and SMS 166
106 Deviation 2: System Reactions in the 14^{th} Quarter 167
107 Deviation 2: Power Plant Cost Distributions 167
108 Deviation 3: SOS and SMS 169
109 Deviation 3: System Reactions in the 4^{th} Quarter 170
110 Deviation 3: System Reactions in the 14^{th} Quarter 170
111 Deviation 3: Power Plant Cost Distributions 171
112 Deviation 4: SOS and SMS 172
113 Deviation 4: System Reactions in the 4^{th} Quarter 173
114 Deviation 4: Power Plant Cost Distributions 173
115 Deviation 5: SOS and SMS 175
116 Deviation 5: System Reactions in the 14^{th} Quarter 176
117 Deviation 5: Power Plant Cost Distributions 176
118 Deviation 6: SOS and SMS 177
119 Deviation 6: System Reactions in the 14^{th} Quarter 178

120 Deviation 6: Power Plant Cost Distributions 178
121 System Total Cost Comparison 179
122 Expansion: Economical Operating Period 180
123 Expansion: SOS and SMS 181
124 Expansion: System Generation vs. Customer Demand 181
125 Expansion: System Cost Distributions 182
126 Histogram of Total LCC 183
127 Scenarios 185
128 Customer Demand Forecasted Under Each Scenario 187
129 System Generation vs. Customer Demand Under Each Scenario 188
130 Scenario 1: SOS and SMS 189
131 Scenario 2: SOS and SMS 189
132 Scenario 3: SOS and SMS 190
133 Scenario 4: SOS and SMS 190
134 Scenario 5: SOS and SMS 191
135 Scenario 6: SOS and SMS 191
136 Scenario 7: SOS and SMS 192
137 Scenario 8: SOS and SMS 192
138 System Total LCC Under Each Scenario 194
139 Scenarios (1-8) 195
140 Scenarios (9-16) 196
141 Customer Demand Forecasted Under Each Scenario (1-8) 197
142 Customer Demand Forecasted Under Each Scenario (9-16) 198
143 System Generation vs. Customer Demand Under Each Scenario (1-8) . . . 199
144 System Generation vs. Customer Demand Under Each Scenario (9-16) . . . 200
145 System Total LCC Under Each Scenario (1-8) 201
146 System Total LCC Under Each Scenario (9-16) 202
147 Scenario 1: SOS and SMS 204
148 Scenario 2: SOS and SMS 204
149 Scenario 3: SOS and SMS 205
150 Scenario 4: SOS and SMS 205

151 Scenario 5: SOS and SMS 206
152 Scenario 6: SOS and SMS 206
153 Scenario 7: SOS and SMS 207
154 Scenario 8: SOS and SMS 207
155 Scenario 9: SOS and SMS 208
156 Scenario 10: SOS and SMS 208
157 Scenario 11: SOS and SMS 209
158 Scenario 12: SOS and SMS 209
159 Scenario 13: SOS and SMS 210
160 Scenario 14: SOS and SMS 210
161 Scenario 15: SOS and SMS 211
162 Scenario 16: SOS and SMS 211
163 Hot-Gas-Path Inspection: Hours-Based Criterion 222
164 Hot-Gas-Path Inspection: Starts-Based Criterion 222
165 Rotor Inspection: Hours-Based Criterion 223
166 Rotor Inspection: Starts-Based Criterion 223
167 Combustor Inspection: Hours-Based Criterion 224
168 Combustor Inspection: Starts-Based Criterion 225

LIST OF ABBREVIATIONS

ANNs Artificial Neural Networks, p. 11.

AR AutoRegressive, p. 101.

ARMA AutoRegression Moving Average, p. 95.

ARMAX AutoRegression Moving Average with External Input, p. 95.

ARX AutoRegression with External Input, p. 132.

CI Combustor Inspection, p. 36.

CWT Continuous Wavelet Transform, p. 83.

DA Decision Action, p. 2.

DCWT Discrete Continuous Wavelet Transform, p. 85.

DM Decision-Making, p. 2.

DWT Discrete Wavelet Transform, p. 85.

EOP Economical Operation Period, p. 65.

EOT Economical Operation Time, p. 154.

FFH Fired Factored Hours, p. 40.

FFS Fired Factored Starts, p. 41.

FFT Fast Fourier Transform, p. 77.

FT Fourier Transform, p. 78.

HGPI Hot-Gas-Path Inspection, p. 36.

LCC Life Cycle Cost, p. 29.

MA Moving Average, p. 132.

MAD Mean Absolute Deviation, p. 148.

MAPE Mean Absolute Percentage Error, p. 148.

MI Major Inspection, p. 36.

MRA Multi-Resolution Analysis, p. 17.

MSE Mean Squared Error, p. 147.

NDWT Non-Decimated Wavelet Transform, p. 89.

OD Optimal Dispatch, p. 2.

PACF	Partial AutoCorrelation Function, p. 101.
SAC	System Available Capacity, p. 65.
SB	Stationary Bootstrap, p. 110.
SCE	System Capacity Expanding, p. 32.
SCEP	System Capacity Expansion Planning, p. 2.
SEEPT	Social, Economical, Environmental, Political, Technological, p. 115.
SMS	System Maintenance Schedules, p. 2.
SOP	System Operation Planning, p. 2.
SOS	System Operating Strategy, p. 5.
SRC	System Reserve Capacity, p. 19.
STFT	Short Time Fourier Transform, p. 78.
TFR	Time Frequency Representation, p. 78.
UC	Unit Commitment, p. 2.
WAW	Wavelet-ARAMX-HoltWinters, p. 97.
WT	Wavelet Transform, p. 81.
XCF	Cross Correlation Function, p. 131.

SUMMARY

In recent years the electric power industry in the United States has been challenged by a high level of uncertainty and volatile changes brought on by deregulation and globalization. This has caused a major restructuring of the electric power industry with the introduction of new and different business practices. The power industry has been split into three entities: power producers, power transmission, and power distribution. This study addresses the power generation part of the overall industry only, and emphasis is placed on power producers that use industrial gas turbine engines even though there are several means for generating electric power. Like any business, a power producer must minimize the life cycle cost while meeting stringent safety and regulatory requirements and fulfilling customer demands for high reliability. Therefore, to achieve true system excellence, a more sophisticated system-level decision-making process with a more accurate forecasting support system is a must to manage diverse and often widely dispersed power generation units as a single, easily scaled and deployed fleet system in order to fully utilize the critical assets of a power producer.

A decision-making process for the fleet management of a power plant has been created as a response to the deregulation of the electric business. A key factor in the process is to take into account the time horizon for each of the major decision actions taken in a power plant and to develop methods for information sharing between them. These decisions are highly interrelated and no optimal operation can be achieved without including information sharing in the overall process.

The decision-making process includes a forecasting system to provide accurate information for planning for uncertainties related to the current power industry. Forecasting has value by offering a better understanding of the forces that might have an impact on the fluctuations in a particular variable, and it improves the quality of decision making by providing a clear picture of uncertainties involved and suggesting contingent strategies.

A new forecasting methodology is proposed, which utilizes a synergy of several modeling techniques properly combined at different time scales of the forecasting objects. It can not only take advantages of the abundant historical data but also take into account the impact of pertinent driving forces from the external business environment to achieve more accurate forecasting results.

By obtaining more accurate information from both the system itself and the external environment, the decision-making process allows for power plants to achieve any-time and any situation system excellence. Then block bootstrap is utilized to measure, based on forecasting information, the bias in the estimate of the expected life cycle cost which will actually be needed to drive the business for a power plant in the long run. Finally, probabilistic scenario analysis is used to apply the proposed forecasting method to realistic situations. The intent is to provide a composite picture of future developments, which may affect the power producer and thus be used as a background for decision making or strategic planning.

To demonstrate an application of the decision-making process, it is applied to a typical (but theoretical) power producer with a certain number of generation units. The power producer was chosen to represent challenging customer demand during high-demand periods. There are limited critical resources for both generation and maintenance to operate the business profitably, and it is necessary to enhance system excellence. The decision-making process proposed in this study achieves this goal by providing more accurate market information, evaluating the impact of external business environment, and considering cross-scale interactions between decision actions. Along with this process, system operation strategies, system maintenance schedules, and system capacity expansion plans that guide the operation of the power plant are optimally identified, and the total life cycle costs are estimated.

CHAPTER I

MOTIVATION

Over the past decade, the electric power industry has witnessed many fundamental and unprecedented changes due to deregulation. This has caused a major restructuring of this industry with the introduction of new and different business practices. The power industry has been split into three entities: (1) power generation at a power plant site – this is the power producer; (2) power transmission from production site to utilities – this is a system that is still evolving and is likely to be controlled by regional transmission operators; and (3) power distribution by utility to customer - this is the power utility. As a result of this realignment, the nature as well as the structure of future electric power industry has become uncertain as the integration of these three systems evolves. This study addresses the power generation part of the overall industry only.

It is recognized that there are several means for generating electric power, including coal fired steam plants, industrial gas turbine engines fired with gas or liquid fuel, hydroelectric, solar, and nuclear power. In this study, power producers that use industrial gas turbine engines are studied exclusively. However, it is hoped that the decision-making (DM) process that is developed will be applicable in part if not in total to any power producing company. The primary objective is to develop an advanced DM process for the power producer operating a fleet of industrial gas turbine engines, which generate electric power for sale in the market place.

An additional aspect of power production as a business is the emergence of companies that offer maintenance contracts to the actual power producer. The most prominent of these companies are the large gas turbine power plant manufacturers such as General Electric, Siemens, and GEC Alstrom, but there are other companies that offer such a service. Thus, a “power producer” in this study refers not only to a company that generates and supplies power, but also to a company that provides power plant maintenance services through contract agreements. These companies, too, are faced with decision making much like that of the power generating company, and the procedures developed in this study, particularly those that pertain to power plant maintenance planning, also apply to these maintenance providing companies.

The uncertainty due to deregulation has necessitated the need for the power producer to determine what is going on, not only within the system itself, but also within the external business environment. To be effective as the industry changes, the power producer must be prepared restructure itself to increase efficiency and reduce life cycle costs (LCCs). This must be done while continuing to satisfy customer demand, which is complicated by unavoidable constraints such as physical operating constraints on the generation units and capacity limits on the total power producer system. This is a tough challenge, and it has put the power generation plant fleet management at center stage in the overall electric power industry, as a power producer depends on its critical assets to operate profitably. Mid- to long-term system maintenance scheduling, operational, and capacity expansion planning for the power producer have received increasing attention in order to enhance system excellence and to achieve the goals listed above under the deregulated environment in which the old rules and regulations are becoming difficult if not impossible to apply. Thus, for power producers, the DM process on the system level has become more important than ever before in history.

Traditionally, managerial decision making in electric power plants has dealt with short-term optimal dispatch (OD), and unit commitment (UC), mid- to long-term system maintenance scheduling (SMS) and system operation planning (SOP), and long-term system capacity expansion planning (SCEP) [36]. Such decision actions (DAs) have different time horizons, which adds another dimension to the DM process and complicates it. The DM process must take into account the time horizon of each of these DAs and identify their appropriate time scales to make consideration of the cross-scale interactions among them possible. Another difficulty emerges when long-term system operating and planning whose time horizons are usually up to more than 10 years become the focus. As the time horizon extends into the future, accurate decision making becomes more problematic due to the increasing uncertainties and complexities that characterize the underlying process, particularly in the current deregulated electric business.

Almost all managerial decisions are based on forecasts because every decision becomes operational at some point in the future. Forecasting data that concerns customer demand, fuel prices, and electricity prices are the main input to the DM process. Such historical data abound in the electric market. Since they are the results of the interactions of many sources that produce different dynamics through drifting and switching, thus providing clues of their development and changes in the past, the main goal of forecasting is to utilize this information to explore the future and support the DM process. Unfortunately, conventional forecasting methods, which create a global model using these historical data, do not recognize these sources and cannot provide satisfactory forecasts. Therefore, a hybrid model architecture that forecasts by accounting for these characteristics should provide high accuracy forecasting results.

The current electric market presents a complex mixture of regulated and deregulated segments. One direct consequence of the transition is rapid changes in the electric market that bring considerable randomness and uncertainty about the future. As a result, the impact of pertinent driving forces from the business environment bears a significant effect on the forecasting process and consequently on the DM process. The forecasting methods solely dependent on the historical data cannot take into account the impact of the external business environment, which generally results in simplified forecasting results. This requires rapidly incorporating external information into the forecasting process and consequently into the DM process.

Thus, a dynamic and adaptive modeling environment and methodology for the power plant system-level DM process must be developed in the current electric market and should consider the following:

1. A multi-timescale DM process that considers the fact that DAs have different time horizons and cross-scale interactions among them.

2. A hybrid forecasting scheme that explores the multi-resolution nature of historical data

and utilizes a synergy of several modeling techniques properly combined at different time-scales.

3. A mechanism for incorporating external information into the forecasting process that takes into account the impact of pertinent driving forces in the business environment.

1.1 Multi-Timescale Decision Making

Decision making plays an essential role in many real world applications ranging from emergency medical treatment in intensive care units to military command and control systems. Existing formalisms and methods have not been effective in applications for which tradeoffs between decision quality and time dependence are essential [6]. In practice, an effective approach to time dependent, dynamical decision making should provide explicit support for dealing with time dependent situations and for modeling their interactions.

The major DAs of an electric power plant can be categorized as short-term OD and UC, mid- to long-term SMS and SOP, and long-term SCEP based on the frequency at which each DA is made and the time horizon during which each DA has an impact. Figure 1 illustrates the time horizons for these DAs and their interactions. The system-level objective of an electric power plant is to meet the customer load and total energy supply demands at any time at a minimum LCC, which requires a tradeoff between responsiveness and efficiency in the operation of the entire system. This requirement must always be considered during each phase of SMS, SOP, and SCEP.

Among the major DAs that have to be taken in a power plant, it is important to mention SCEP [78]. It is the study of determining the generating resources required to meet the growth in demand at the lowest possible cost in a long run, considering environmental and financial constraints. For example, SCEP is needed to determine what generation units should be constructed and when they should come online over a long-term planning horizon [50]. Since a power plant should meet customer demand under a wide range of normal, abnormal, and emergency conditions, including foreseeable maintenance outages and unforeseeable failures of facilities, it must have capacity reserve in excess of the forecasted

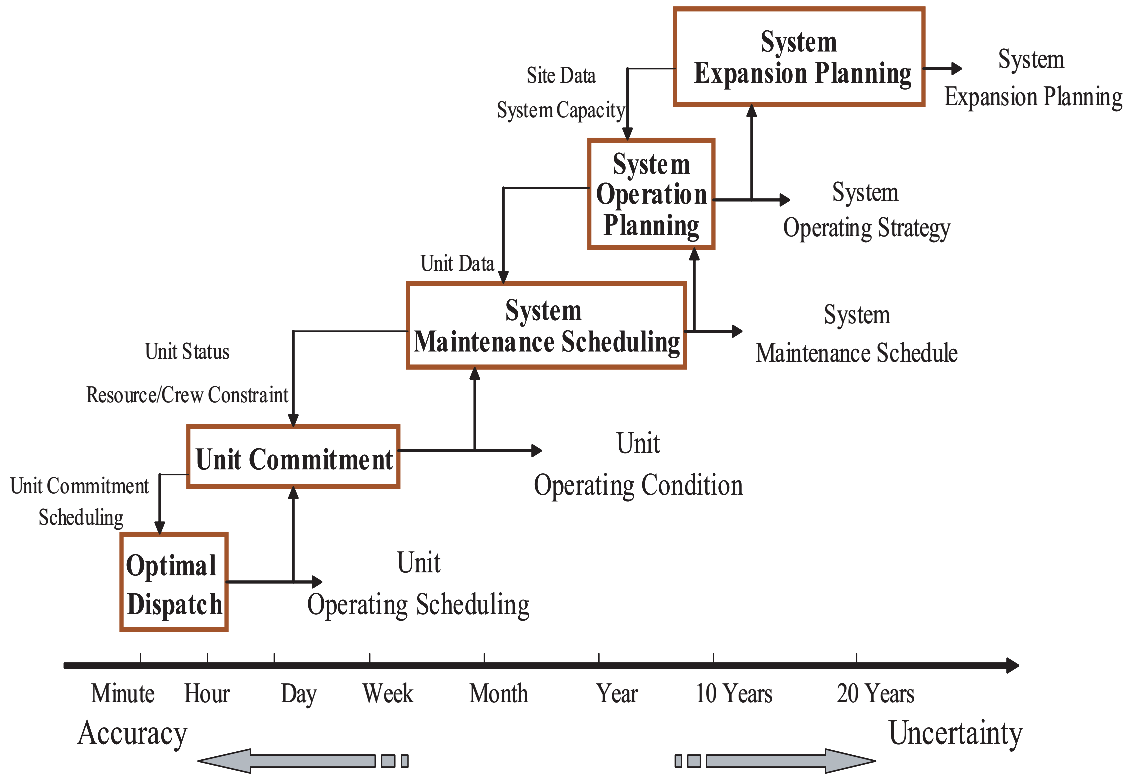


Figure 1: Time Scales of Major Decision Actions

customer demand [10]. In a competitive environment, decision makers who are considering alternatives for an expansion of generating capacity have to consider various sources of uncertainties resulting from the remote future target and the volatile business environment.

During the past, efforts have largely concentrated on decision making in SCEP, but nowadays, they must be extended to power plant operation. Long-term SOP is defined as the process of evaluating alternative system operation strategies (SOS) against the desired objectives subject to technical, environmental, and contractual requirements and selecting a recommended strategy with a time horizon that extends beyond one that requires immediate commitments. The need for a long-term system operation has become more pressing than in the past because of the rapid changes in economy, fuel resources, environmental constraints, and so forth [91]. Uncertainty must also be resolved by through changes in the time horizon from system operation to system capacity expansion.

An area related to the system operation is decision making in SMS, including maintenance types and their effects, maintenance optimization, spares policy, and residual life studies. In addition, assessment of risks to power plants is required due to the different maintenance schedules they adopt. Maintenance scheduling at the system level makes it easier to satisfy customer demand both responsively and efficiently, especially when unscheduled events (e.g., unscheduled maintenances) take place. From a system point of view, the determination of the sequence and the maximum number of units that should be taken offline for scheduled maintenance must consider the following constraints:

- The maintenance resources that a power plant owns. All the units have to share the limited resources of the system, which includes employees and material inventory. They determine the maximum number of units that can be under maintenance simultaneously from a resources point of view. If the number exceeds the limit, maintenance delay and its related costs have to be taken into account as it influences system LCCs and diverts the system from the optimal operating condition.

- The generation resources that a power plant owns. The power plant must satisfy customer demand at any time. Offline generation units might contribute to system generation. So the generation resources determine the maximum number of units that can be under maintenance simultaneously from a demand point of view. They also determine the sequence that each individual unit can be taken offline. For example, the units that contribute significantly to customer demand should not be taken offline during high-demand periods unless it is absolutely necessary.

The enormous complexities in the evolution and revolution of power plants render decision making difficult. Making the situation even more complicated, many of the dynamics occur at vastly different time scales and are influenced by cross-scale interactions (see Figure 1). The time scale of a generation unit in maintenance is many orders of magnitude shorter than the dynamics of that generation unit in operation. The former takes place on the order of several hours to several weeks, while the latter may have a time constant of years. Compared to the operation process, maintenance can be treated as a "point event" or a "short-term event" that normally disrupts the comparatively long-term operation process.

Cross-scale interactions result when the events at one level of scale influence events at other levels. Scale variations have long been known to constrain the details with which information can be observed, represented, analyzed, and communicated across scales. A "one size fits all" approach to the assessment of cross-scale interactive problems can result in problems involving scale mismatch or ignorance of cross-scale linkages. Problems also exist when two or more assessments of the same issue done at different levels of scale compete for the attention of overlapping audiences. Decision makers increasingly recognize the importance of scale and cross-scale dynamics. Time scale decomposition is a way to achieve enhanced information sharing between scales and better deterministic model approximation. Thus, synthesizing current practices and theories about scale and cross-scale interactions in the DM process has become one focus of this study.

As a result, a DM process that takes time scales and cross-scale interactions into account is needed to overcome the intractable nature of an exact centralized DM problem, exploit information sharing, and capture the essence of the stochastic dynamics of it.

1.2 The Forecasting Problem for Power Plants

So-called "forecasting" is a process in which one studies given objects or affairs (forecasting variables) to find clues of their development and changes in the past, explores rules of their development using scientific tools such as statistical methods and systematic identification, and finally makes estimates for the future changes [112]. Clues may include the historical data of the forecasting variables and the historical or forecasting data of other related variables. Forecasting is a dynamic and continuous process affected by the development of different fields and other factors. It can be divided into three steps: forecasting problem identification, forecasting problem solving, and assisting with decision making.

- **Forecasting Problem Identification**
  - Identify the forecasting objective and forecasting environment
  - Analyze forecasting variables (forecasting objects) and related factors (forecasting environment) and their relationships; for example, if customer demand of electricity is the forecasting variable, the forecasting environment is the electric market.

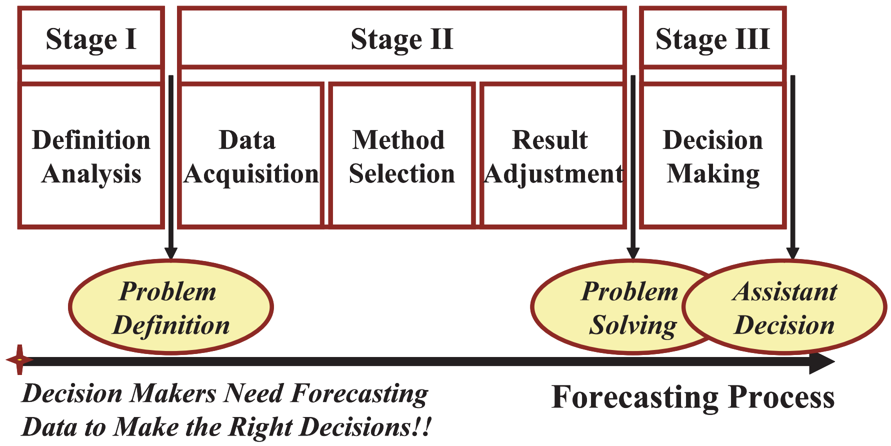


Figure 2: Forecasting Process

- **Forecasting Problem Solving**
  - Obtain source data for forecasting variables and related factors
  - Explore the data characteristics
  - Select a forecasting methodology
  - Perform forecasting
  - Analyze and adjust the forecasting results according to expert experience
- **Assisting with Decision Making**

  Utilize forecasting results to assist the decision makers

Forecasting objects and the environment refer to the forecasting variables and related factors, respectively. Compared to the identification of forecasting objects, the identification of the forecasting environment, which has an impact on the development of the forecasting objects, is more difficult. Using unrelated factors or neglecting important related factors will result in lower accuracy in the forecasting. In the next section, the forecasting environment will be discussed in detail. The whole process of forecasting and the tasks in each stage are illustrated in Figure 2.

In addition to forecasting, there are two other important factors must be considered: forecasters and forecasting technology. Forecasters, or decision makers, are the subjective factors in the forecasting system. Forecasting technology includes various forecasting models and methods. These three elements together constitute the forecasting system.

The forecasting system, which can be in either implicit or explicit form, is a sub-system and support system in the DM process. Electric power plants require accurate forecasting data at each step of the DM process because they need to ***plan*** for an ***uncertain*** future [97]. Therefore, most power plants operators spend considerable time and effort in forecasting.

Luck plays some role in business success, but not a strong role. More crucial to the business success of power plants is the careful selection of system operating strategies and maintenance schedules, and allocation of precious resources to satisfy customer demand efficiently and responsively. To achieve these goals at any time in the future, decision makers must determine the appropriate actions to take. They can accomplish these goals by comparing actual system operational behaviors and properties with the estimated ones and then making necessary changes in maintenance and operation so that they can be implemented accurately and efficiently.

Another strong reason to develop good forecasting technologies for power plants is that the future is uncertain. Uncertainty is involved in nearly all analyses of electric power plants. For example, uncertainties have always existed, to some degree, in customer demand, fuel prices, electricity prices, and capital costs. However, additional uncertainties resulting from deregulation and restructuring are now further complicating the DM process of power plants. Thus, forecasting has value by offering a better understanding of the forces that might have an impact on the fluctuations of a particular variable and improves the quality of decision making by providing a clear picture of the uncertainties involved and suggesting contingent strategies.

Any inaccuracy in forecasting significantly affects the DM process. For example, inaccuracy in customer demand forecasting may result in overbuilding of supply facilities and unprofitable operation in cases of overly optimistic forecasting, or it might curtail customer demand and cause poor system reliability in the cases of overly pessimistic forecasting. Both cases are unacceptable because they affect profitability [69]. In the latter situation, the penalty for not supplying customer demand is very high in the deregulated electricity market. For example, in Australia, the cost of loss-of-load is AUD$5000 per KWh (valid as up to the year 2000 [7]). Therefore, a power plant may lose a whole year's revenue due to an unexpected loss of generation caused by some contingencies [111]. Also strategically important to power plants is reliable fuel price forecasting data. In most power plants, fuel accounts for 60% to 80% of operating costs, and for 20% to 40% of the total cost of electricity [13]. Fuel expenditures are typically hundreds of millions of dollars a year. Therefore, any inaccuracy associated with fuel prices and capital costs would profoundly affect a power plant once the generation units are built rendering a sound economic decision a poor one because of the time lag of several years between the decision and completion of the DA. In extreme cases, such a situation would result in significant financial hardship to a power plant [10].

One of the most useful criteria for matching a specific forecasting situation with the most appropriate technique is the time horizon. Since mid- to long-term system planning, the focus of this study, is crucial to power plants, forecasts should be provided on annual, quarterly, or even monthly basis depending on the actual forecasting horizon. Forecasting results should support scheduling maintenance, planning operation, and future capacity expansion in order to determine the level and direction of cost expenditures. It is in the field of strategic planning that the greatest value of forecasting lies.

From both a theoretical and a practical standpoint, forecasting for mid- to long-term planning is radically different from that for short-term planning, and therefore, it necessitates different treatment. An important characteristic of long-term situations is that the time lag between the point at which a forecast must be performed and the actual occurrence of events is quite long. The uncertainty associated with the forecasting increases as the time horizon elongates into the far future because the future is never exactly the same as the past. That means the confidence limits of establishing accurate forecasting broaden as the time frame of forecasting increases, reflecting a growing level of uncertainty. An analogy of this type of uncertainty is the forecasting of the price of fuel (e.g., natural gas). One could forecast the price of fuel on the next day with a very high confidence. However, forecasting the price in the next 20 years would yield very low confidence [51]. Few people could or did, for example, foresee the decline in the growth of the railroad of the last few decades or the saturation in sales of the glass and aluminum industries. Hence, the decision making based on forecasts over the long-term time horizon involves higher levels of uncertainty and volatility. On the one hand, this creates a need to facilitate accurate data analysis in order to reveal the underlying driving forces that result in fluctuations in the forecasting variables. On the other hand, the forecasting process should never end but instead should be updated periodically, as the time of certain events approaches, or as more information relevant to that situation is obtained, as well as the decisions that follow.

Current techniques for forecasting can be broadly classified into two groups: factor analysis methods and time series methods. Factor analysis methods, also known as causal methods, are based on the determination of various related factors that influence the forecasting variable. Their correlations with the forecasting variable are calculated to discover the form of the cause and effect relationship that will be used to forecast future values of the forecasting variables. Time series methods forecast the future based on the historical data of the forecasting variable by discovering its underlying pattern and extrapolating that pattern into the future. In recent years, artificial neural networks (ANNs) have demonstrated an impressive ability to deal with forecasting events when the networks have a large database of prior examples to draw on. Based on these approaches, a number of different methods have been developed in the past, literature [23], [84], [65], and [3] provide an overview of some of the commonly used methods. These above mentioned conventional models fail to give reasonably accurate forecasts for electric business-related problems because of their inherent limitations. Factor analysis methods are inefficient as forecasting of the related factors itself is not easy, and time series methods are not adaptive to sudden changes that last a short period of time. The implementation of ANNs still suffers from a lack of efficient constructive methods for both slow convergence and the determination of the network structure and parameters [114].

Advanced techniques that accomplish the task of forecasting in the electric business should be utilized instead. An important prerequisite for the successful application of some modern advanced forecasting techniques is a certain uniformity of the forecasting variables

[64]. Generally, historical data such as customer demand, natural gas prices, and electricity prices contain a very wide range of frequencies and harmonics from the extremely long wavelengths such as trends to very high frequency transients caused by short-term special events. The existence of different kinds of non-stationarities is due to the fact that these data series may be the result of the impact of various forces that drift and interact, producing different dynamics. Conventional approaches usually provide one best or global model that characterizes the measured historical data. When a data series is non-stationary, as is the case for most time series in electric markets, identifying a proper global model becomes extremely difficult.

To facilitate accurate data analysis and to reveal aspects that global model techniques miss, a robust high frequency filtering, seasonality identification, and trend analysis method must be utilized as it affords a different view of the data than that provided by conventional techniques. The most efficient way is to design a hybrid scheme, ant then to utilize a synergy of several modeling techniques properly combined at different time-scales through multi-resolution analysis techniques such as wavelet transforms. Wavelet transform can analyze data at different frequencies with different resolutions, and thus produce a good local representation of them. Unlike the Fourier basis, wavelets can be supported on an arbitrarily small closed interval. Thus, wavelet transform is a very powerful tool for capturing transient phenomena that are taking place in the current electric market. Combining wavelet transforms in the historical data analysis and a hybrid forecasting scheme can provide better forecasting results for the electric business.

1.3 External Information Adaptive Processing

The electric power industry has traditionally been a regulated monopoly that was structured in a single vertically centralized, integrated organization for providing electric power to its customers. In the restructured market, information structure and the DM process have become more decentralized and more distributed. At the same time, another great revolution is taking place, that of the strength of information replacing mechanical strength. The heart of this transition is that it is primarily information that provides an economic advantage, but not necessarily the physical scale. The deregulation of the electric power industry, coupled with the emergency of an information-based economy, is a double engine that makes this the most energized time ever to be in the electric business.

Information is rapidly becoming the key to profitability, customer retention, market advantage, and business growth in the increasingly competitive electric power industry [67]. The electric power industry requires information input not only to perform traditional real time functions for operational and commercial purposes, but also to support the new functionality that specifically meets the needs of competition and uncertainty resulting from deregulation. As a result, power plants want more than ever to use information in innovative ways to improve forecasts and consequently to improve the quality of decision making in order to lower LCCs, improve customer satisfaction, and increase market share to enhance system excellence. This requires a comprehensive system that enables communication and integration of external information in the DM process in the electric power industry.

The most important source of information comes from the business environment. Very little in the business environment is stable and unchangeable. In fact, almost all business and industries have fluctuating patterns. The key to success is not to wait until these trends hit one hard but instead to identify any precursors, so that appropriate actions can be taken to soften the impact. For example, many socio-economic activities and natural causes directly affect the forecasting process and the development of power plants. A non-exhaustive list includes the following:

- Seasonal variations, e.g., customer demand is a function of time of month, week, or even day.
- Weather, e.g., extremely low or high temperature is responsible for increased heating and air conditioning load, respectively.
- Special events, such as major sports gatherings, system outages, Severe Acute Respiratory Syndrome (SARS) outbreak in Asia in 2003, and other events.
- Known future events, such as public holidays.

- Economic growth, changes in the employment rate, GDP, and so forth.

The impact of these socio-economic conditions, climatic conditions, and special events on the rapidly changing electric business becomes stronger especially for long-term system operation and planning during periods of deregulation. For DM problems over a short-term horizon, the current operating states of power plants should play an important role, whereas in the long-term operation and planning, the static state condition can be assumed only if the changes in the environment and the system itself are so small and slow that their effects can be neglected. The deregulation and restructure of the electric power industry result in a high level of uncertainty and randomness in the future for each hour of the study time horizon, which has created a need to change the way such information is processed and decisions are made.

Adding to the complexity is the fact that some of these changes in the business environment occur rapidly, and their effects may disappear in a short time. For example, extremely hot weather in the summer definitely influences demand, but for only a certain period of time. This phenomenon may not be captured in the use of a purely causal forecasting model. However, the likelihood that the forecasting methods utilizing trends would reduce this effect is small, because most forecasting methods that utilize trends are not local (in terms of time) in nature. For example, it is counter intuitive to conclude that the effects of the SARS outbreak in Asia in early 2003 will still be experienced five years later, which would be suggested if traditional trending methods were used.

Therefore, the results provided by the forecasting methods that do not take into account the possibility of changes in the external business environment and that depend solely on intrinsic historical data are generally too simple. The last section also mentioned that missing important related factors will result in a lower accuracy in the forecasting results. This implies that new methodologies, models, and technologies that reflect changes in the business environment and new concepts for system planning and decision making that cope with the new circumstances must be proposed [115]. The key problem is how to incorporate pertinent external information immediately into the forecasting process and consequently into the DM process.

One important characteristic of information is that it has time value. The value of information or data several years ago should be less than or equal to the value of the data or information collected today. The present value of historical data or information is a function of how old they are, and what has happened since the data or information were collected [83]. As time elapses, the system is less responsive to information collected in the past. The electric business is a rapidly changing industry, so incorporating new information into the system as soon as possible is a sensible measure to take. Figure 3 shows that the value of information is perishable. As time passes, given information changes from an operational status to a decision support status and finally to a historical status, which is called statutory or "shelf-life" status.

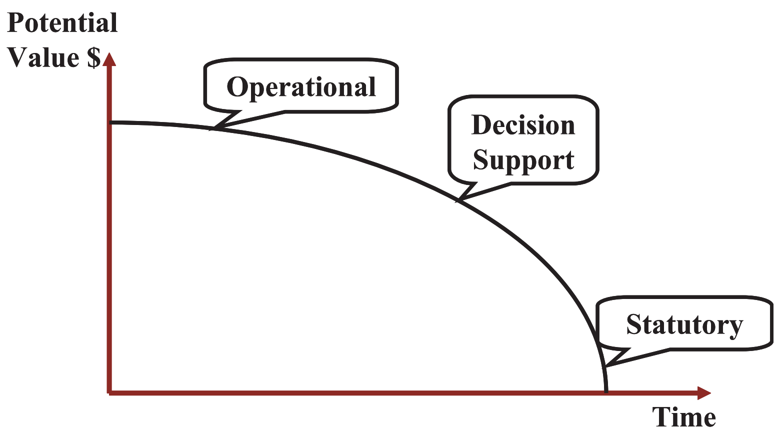


Figure 3: Time Value of Information

In summary, the high volatility of the electric market complicates long-term decision making for power plants. Thus, an adaptive modeling tool is needed that has a mechanism that incorporates external information with a short lead time such that it can update the estimates after each new observation is obtained and utilize the information in the next step of the DM process. Additionally, communicating with the external business environment and integrating it into the DM process are excellent ways of preparing decision makers to face the uncertainties of the future and help them realize the potential impact of some key driving forces that may influence the future development of the power plant. Therefore, a systematic and consistent treatment of the various sources of information must become an integral part of the DM process. However, this requires separate consideration by either pursuing the search for more improvement in the existing forecasting techniques or establishing another approach that addresses this problem.

1.4 Research Questions and Assumptions

The challenge of the current research is to formulate a physics-based, system-level DM process that can help power plants reduce life-cycle costs and satisfy customer demand through improvements in the forecasting methodology and the DM process. Accomplishing this goal requires that cross-scale interactions be addressed, hybrid scheme of forecasting be utilized, and external driving forces from the business environment be incorporated into the forecasting process and consequently into the DM process. The DM process must be able to capture the real optimal system operating strategies and suitable system maintenance schedules that will produce enough power to satisfy customer demand under any circumstances at a minimum LCC, which includes maintenance and operating costs for the existing power plant and investment, maintenance, and operating costs for the capacity expansion planning. This process must be able to determine the optimal number and time for the introduction of new generating units for expansion planning. This process must have a forecasting capability as a support system whose accuracy and subsequent accuracy in estimating the expected LCC must be identified. This process must investigate how the system will develop under different external environments due to the uncertainty involved throughout the process.

1.4.1 Research Questions

The identification of these needs leads to a multitude of research questions that this study will attempt to resolve. These research questions are as follows:

The research questions that must be addressed to consider the multi-scale DM problem include:

- *How will the cross-scale interactions be accounted for?*
- *How will the timescale for each DA be determined?*
- *How will "point events" be handled?*

The research questions that must be considered when developing the forecasting method include:

- *How will data analysis be facilitated by utilizing MRA (non-decimate wavelet transform) to extract critical information from historical data for forecasting?*
- *What available modeling techniques can be appropriately applied to each time scale? How will external information be incorporated into the forecasting process?*
- *How will the behavior of forecasting errors be identified?*

The research questions that must be considered in the evaluation of the impact of the external business environment are as follows:

- *How will the bias of the estimate of the LCC needed to drive the business be evaluated?*
- *What are the critical sources of uncertainty and their features?*
- *How will the uncertainty from the external business environment be explored?*

1.4.2 Hypotheses

The loss of production due to non-perfect maintenance and performance degradation is assumed small when compared to the loss of production during generation contingencies. Each generation unit can be treated as a "black box" with inputs and outputs available.

The impact of external driving forces on the power plants can also be evaluated. The use of statistical and probability theories will enable the quantification of their impact and the exploration of the evolution of power plants, which will provide subjectivity to the DM process.

CHAPTER II

BACKGROUND

Before the research questions addressed in the last section of chapter 1 are answered, the major DAs of electric power plants and their interactions are reviewed. In addition, the definitions and methods that could be used in the development process are identified. With regard to the forecasting system required for the DM process, the current electric market, the identified forecasting variables, and the forecasting techniques that are currently in use and their deficiencies are addressed.

2.1 Major Power Plant Decision Actions

DAs for power plants are usually arranged according to their time horizons, which usually consist of two main levels: long term and short term. Long-term DAs usually have a study time horizon of more than five years and essentially include maintenance and fuel resource scheduling, operation planning, and capacity expansion. The short-term DAs include UC (usually a week), and OD (from one hour down to a few minutes). The different time horizons provide a typical hierarchical planning structure (see Figure 1). However, the interactions between these DAs complicate the picture. The objective, inputs, and outputs for each DA are discussed in detail in the following sections.

2.1.1 Optimal Dispatch

An operational planning in the electric power industry concerns the operational strategy that a power plant adopts to operate its generation units. For short-term operational planning, two major DAs are considered [62]. One is the UC, which determines on/off schedules of generation units in order to minimize the overall system operation cost over the planning time horizon, and at the same time, satisfy customer demand and meet system constraints. To complete short-term operational planning, another problem that needs to be resolved is how to determine the assignment of generation power for each committed unit to minimize fuel costs without violating the unit's generation limits. This DA is usually called "optimal dispatch." In order to achieve the optimal operation condition, UC and OD must be simultaneously performed.

The daily operation of a fleet of geographically dispersed generation units entails the well known problem of OD. The concept of OD, which first appeared in the 1950's and was used by the electric power industry [82], is defined as the process of allocating generation levels to a fleet of dispatchable generation units so that the required power is produced while minimizing the fuel cost of generating real power, and that a minimum system reserve capacity is provided over a given period of time, from 15 minutes up to 24 hours. Based on customer demand forecasting and the specific properties of a power plant, the optimal operation schedule has to be determined. It affects not only the economic interests of a power plant, but also the stable and secure operation of the power plant [22].

Minimum reserve capacity is operationally required to ensure a sufficient reserve so that the power plant can respond within a specified time to a generation contingency and/or a demand contingency. Generation contingency is caused by the loss of a single generation unit. Demand contingency occurs because of the unexpected increase in customer demand. These two contingencies are different and should be treated as such. A certain amount of system reserved capacity (SRC) for demand contingencies can be determined through OD, but if a generation contingency occurs, reserve capacity actually available to remedy it depends on the contingency itself. Hence, complete certainty about the amount of available SRC is not possible since the outaged unit may have contributed to the reserve requirement.

From the economic aspect, OD concerns how to achieve the minimal fuel cost by distributing customer demand over a fleet of dispatchable units. Fuel cost is a major component of a power plant's LCC, which generally includes fuel costs, emission costs, operation and maintenance costs, and network loss costs. Reducing fuel costs by as little as 0.5% can result in enormous annual savings. Therefore, the economic consequences of OD are crucial.

An OD problem falls under the class of a constrained optimization problem [103] with the objective of minimizing operational costs and constraints as listed below:

- Downward and upward regulating margin requirements of the system.

- Lower and upper economic limits of each generating unit.
- Maximum ramping rate of each generation unit.
- Unit's restricted operating zones.
- Emission allowance of the system (so_2, co_2, no_x).
- Network security constraints (maximum MW power flows of transmission lines).
- Supporting multiple I/O curves (incremental heat rate) and emission cost curves for different fuels.

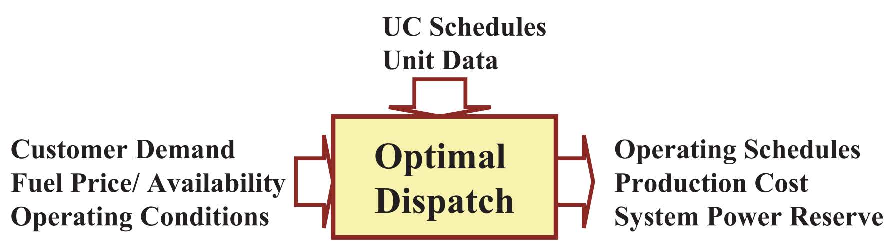


Figure 4: Optimal Dispatch Decision Action

Figure 4 illustrates the input and output variables of this DA. Customer demand is a major input provided through the forecasting system. Fuel requirements include fuel price and availability. Fuel price can be obtained from fuel price forecasting. Fuel availability is usually ensured through long-term contracts. Unit data include the operating states of each individual generation unit. The operating conditions of the system are determined, therefore, from the individual unit data. The UC schedule is the feedback information from the UC problem. The outputs include the operating schedule, which allocates the customer requirements among the available generation units, the production cost, which is the minimum cost that can be achieved, and the system power reserve, which measures the reliability of the system.

2.1.2 Unit Commitment

As is true for many systems, electric power plants experience different cycles. Customer demand in one day is higher during the daytime and lower during the late evening and early morning. This cyclical demand requires that power plants plan for producing power on an hourly basis. The first problem is deciding the units to be turned on, UC, and then determining an OD schedule for these dispatchable units so that they meet customer demand while satisfying the operational constraints. The purpose of UC is to plan for making sufficient available generation units that meet customer demand the next day or the next week.

UC is a very important DA for the economical operation and short-term planning of power plants. The objective of the UC problem is to determine a minimal cost turn-on and turn-off schedule of a set of generation units that meet customer demand while satisfying the operational constraints [86] for a given period of time. It is a nonlinear, large-scale, mixed-integer combinational optimization problem [81]. The optimal solution to this problem leads to remarkable savings in the cost of system operation. However, this solution is quite complex because of the enormous dimension, the nonlinear objective function, and the large number of constraints [2]. The exact solution can be obtained by a complete enumeration of all feasible combinations of generation units, which could be considerable.

Many constraints need to be imposed on UC on the system level. At the same time, each individual generation unit may specify its own set of constraints, depending on its own properties, such as load curve characteristics, and reliability and security requirements. Thus, the constraints can be classified into the following two categories:

- **System constraints**. System constraints are applied to the objective function from the system level in order to keep the power plant within the acceptable stability and security limits. The most common system constraints are listed as follows:

  - The total generated power must be equal to the demand.
  - Sufficient system reserve power must be available in cases of demand contingencies or generation contingencies, or both. System spinning reserve is defined as the extra amount of power that can be obtained from the committed units within a specified period of time, e.g., a few minutes by loading them to their maximum rating.

- **Unit constraints**. Unit constraints are applied to the operation of each individual generation unit and vary from one unit to another. The most common unit constraints are as follows:

  - The production by each unit must be within certain limits (minimum and maximum capacity).
  - Minimum uptime t_{up} states that a unit that is running must be up for at least t_{up} hours. The uptime constraints arise from physical considerations associated with thermal stress on the units and are designed to prevent equipment fatigue.
  - Minimum down time t_{down} states that a unit that is down must stay down for at least t_{down} hours. Minimum downtime constraints are based on economic considerations intended to prevent excessive maintenance and repair costs due to frequent unit cycling.
  - Loading and de-loading rate of the unit.
  - Must off units.
  - Must run units.
  - Crew constraints.

The total operation cost in the objective function includes two major terms. The first is fuel costs, or the cost of producing the power required, which depends on the amount of fuel consumed and the fuel price per unit production. The second is the start-up cost, which depends on the prevailing temperature of the generation units. The total cost of the online units can be obtained by adding these two costs. However, the total cost can be minimized by the proper manipulation of some variables, subject to the necessary constraints.

The start-up cost, which relates to turning a unit on, is determined by one of the following two types of start-ups: a cold start-up cost, which will be incurred if a unit has been off for a long period and the temperature of the equipment becomes close to the ambient temperature, and a hot start-up, which is applied if a unit has been recently turned off and its temperature is still close to the normal operating temperature [102]. Therefore, the start-up cost is a function of the period of time for which the unit stays down. Its value may vary from the maximum value for a cold start-up to the minimum value for a hot start-up.

Fuel costs represent a significant part of the total operation cost and is a function of unit efficiency, and therefore, they will significantly be affected by the selection of units that meet the forecasted customer demand. An increase in demand requires that the most efficient available unit in the system be put into service. When the demand declines, less efficient units would be taken off line first. As a result, the lowest possible cost can be achieved by the appropriate selection of units, taking into account system and unit constraints. This process is performed at least once a day to cover a period of twenty-four hours. It may be extended over a longer period, perhaps a week or ten days in advance.

Figure 5: Unit Commitment Decision Action

Figure 5 shows the input and output variables of this DA. Input variables include customer demand, fuel price/availability, unit data, and operating conditions. The system and unit material resources and crew resources act as constraints for the UC problem. Output variables include the production cost, system power reserve, and UC schedules, the last of which are an input to the OD problem.

2.1.3 System Maintenance Scheduling

Due to the critical importance of electric energy and the rising cost of its production, power plants are compelled to minimize production costs as well as hidden costs for failing to meet customer demand and for introducing new units to increase system capacity while operating with sufficient reserve to ensure an acceptable level of system reliability. The efficient operation of electric power plants requires the solution of several inter-related problems. One problem that has proven to be particularly unyielding is that of determining when each generation unit should be taken out of service for scheduled maintenance, or preventive maintenance.

Scheduled maintenance of generation units, an important part in the overall power plant management, has attracted enormous attention of planners and designers in the electric power industry. A SMS, a regular routine of planned checkups and repair over a one or two-year operational planning period of a fleet generation units, is required to reduce the probability of capacity shortage, to improve the overall system reliability of power plants, and to minimize the total operating cost while satisfying maintenance constraints. In detail, SMS specifies the periods of the operation process during which each generation unit is to be taken off line for scheduled maintenance while considering forecasted customer demand, and the maintenance requirements and constraints. Because units under maintenance are not available to the system, the total installed capacity decreases, contributing to lower system reliability and higher production costs. Scheduling maintenance should, therefore, take into account both system reliability and production costs. Correspondingly, energy costs can be divided into two parts: energy production costs and reliability costs.

From the point of view of system reliability, all power plants perform scheduled maintenance in order to ensure that the equipment is always in operation, to reduce equipment faults, to extend equipment life, to reduce frequency of service interruptions, and therefore, to increase reliability. Whichever maintenance schedule is employed, a selected unit has to be taken out of service for periods of time ranging from several hours to several weeks. Usually, unit outages have a detrimental effect on overall system reliability, which can range from negligible to significant, depending on the load carried and the degree of redundancy available. As a result, maintenance is usually performed at the most suitable time from the system reliability point of view [88]. A good SMS can improve the reliability of the system and balance customer demand among different areas.

Suboptimal SMS not only contributes to lowering system reliability but also increasing system production costs by adversely affecting many short and long-term operations planning functions [63] such as UC, fuel scheduling, and SOP, all of which have maintenance schedules as an input. For example, schedules with high reliability tend to have high production costs, and vice versa. However, a schedule that provides the highest reliability may not have the highest production cost. Hence, optimizing SMSs is significantly beneficial. The optimal solution among the many feasible schedules is one that minimizes the operational cost over the operational planning period subject to unit and system maintenance constraints [113].

Because SMS plays a very significant role in the economical and reliable operation of power plants, the following methods have been applied in an effort to solve this problem:

- The classical approach is based on leveling the reserve throughout a period of time. This approach has been widely used because of its simplicity. The main drawback is that it is deterministic in the sense that uncertainties, for example, the uncertainty involved in customer demand forecasting, is not taken into account.

- The approach based on leveling the system energy costs attempts to minimize the unit maintenance cost [55]. This approach considers both production costs and reliability. If a unit is put under maintenance too early, a part of the investment made during the previous maintenance is foregone, as it was meant for a longer duration of operation of the unit. On the other hand, deferring maintenance of a unit beyond the maximum period involves extra expenses for maintenance caused by partial or full damage of the unit. This method seeks a trade-off between the two.

Past studies [88] have shown that schedules that are optimized in terms of one criterion are usually quite good compared with others. In particular, leveling net reserve does not lead to much riskier schedules. The two parameters, namely reliability and production costs, are both important in decisions regarding the maintenance schedules of generation units. Therefore, maintenance problems have always been investigated together with system reliability problems.

With the appearance of the deregulation and the restructure of the electric power industry, SMS has acquired a number of new features that differentiate it from that in the traditional centralized electric power industry [98]. Traditionally, a meta-system dispatch center coordinates the various maintenance schedules of all the power plants with respect to an optimization objective of the meta-system, such as leveling either the energy reserve rate or the risk rate. At the same time, it tries to ensure that units within one region are not placed under maintenance simultaneously so that energy supply is sufficient and energy transmission secure within the meta-system. In the deregulated electric power industry, unit maintenance schedules will no longer be coordinated by the meta-system dispatch center. Instead, power plants' decision makers will coordinate their own maintenance schedule without considering the maintenance schedules of other power plants. They will schedule maintenance according to the operating conditions of their units, the quotations on the energy market, and other economic factors [108]. Their goals of this approach are to extend the life span of their units and to maximize the profit from their production.

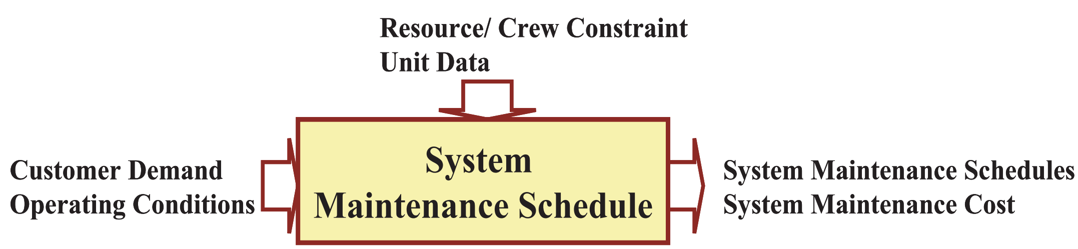


Figure 6: System Maintenance Scheduling Decision Action

Figure 6 illustrates the input variables, output variables, and constraints for this DA. The input variables include customer demand forecasting data, unit data, and system operating conditions. The constraints for this DA include unit and system material resources and crew resources, the maximum number of units that can be under maintenance simultaneously, and a maintenance window, which is the continuous time frame within which the maintenance activity should be completed. The output variables are the SMSs and costs.

2.1.4 System Operational Planning

In the context of power plants, long-term SOP is basically a study of how a power plant should be operated at some time in the future. It includes the study of determining which generation units should be committed and what level of load should be placed on each such that forecasted demand is met over a period of time that is beyond the immediate UC.

As mentioned above, power plants experience different cycles. UC deals with the daily cycles in customer demand. SOP usually covers a period of several months to several years in the future and is used to exploit the flexibilities of power plants when dealing with seasonal cycles in customer demand. Commonly, customer demand in a year is higher during the summer, the early fall, and the winter and lower during the spring and the late fall. The profiles for different seasons are different [66] (see Figure 7). It is highly desirable that power plants plan for production considering the seasonal characteristics of customer demand.

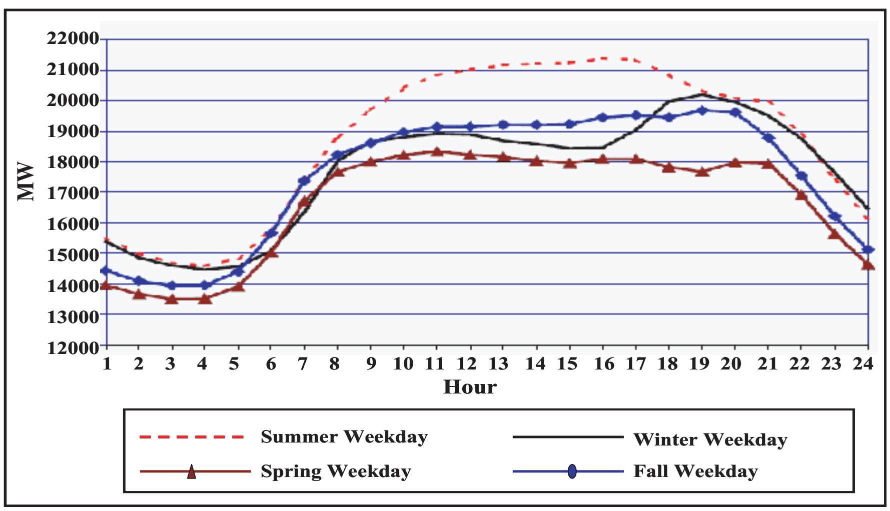


Figure 7: Load Profiles for Four Seasons

A wide variety of research chiefly focusing on the analysis of the commitment decisions from the short-term perspective has been done. SOP received much less attention than short-term UC, partly because the problem is much more complex. As the electric power industry is moving away from regulated monopolies and toward a more uncertain, competitive environment, mid- to long-term power plant operational planning is awakening the interest of researchers and is becoming a subject of importance.

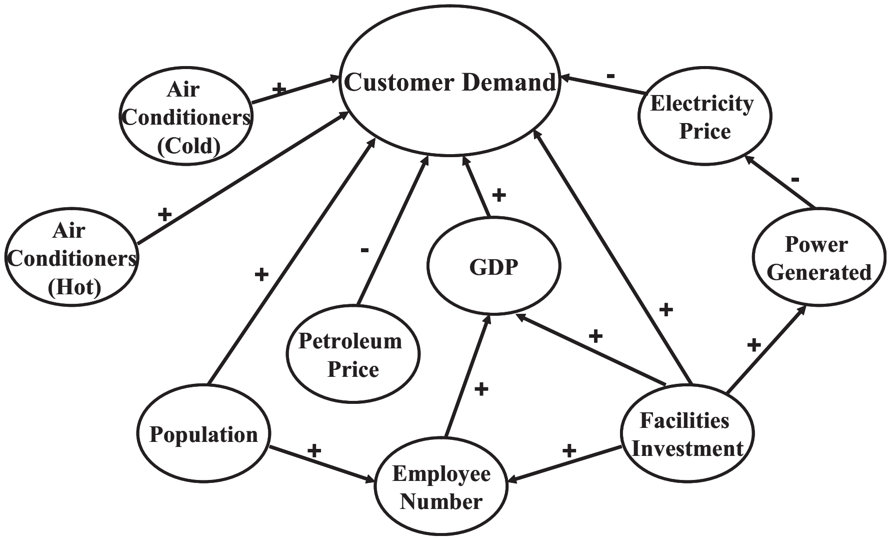


Figure 8: Contributing Factors to the Volatility of Customer Demands

Even though the power plant itself will not change appreciably due to the introduction of new units or the scrapping of old units during the planning period, SOP has to consider the following factors, all of which contribute to the extra difficulties in the DM process:

- Customer demand forecasts become less accurate. Despite the important role that long-term customer demand forecasting plays in power plants for SOP, it is inaccurate because it is affected directly or indirectly by various related factors, illustrated in Figure 8 (not an exhaustive list). Satisfactory forecasting can be achieved by taking into account these definite and indefinite relations. However, as the study time horizon extends into the future, it becomes less possible to consider the interactions between the customer demand and these related factors due to uncertainties involved in this process [46].

- Fuel resources are less determinate. For most power plants, fuel costs are the largest single operating cost component. As fuel availability becomes less determinate, fuel price becomes more variable due to the impact of related factors such as transportation limitations, storage costs and constraints, environmental, socioeconomic, and political considerations, contractual obligations, and other such factors. They render the fuel scheduling problem a major concern in the long-term SOP for many power plants.

- Cross-scale interactions are involved. When SOP is being considered, one problem that cannot be ignored is the SMS problem. As discussed in Chapter 1, these two DAs are closely inter-related. Maintenance activities aim at operating the system with a high level of reliability and security. However, the generation units under maintenance might contribute to lower SRC and higher production costs, leading to a tradeoff between how to appropriately commit and operate the generation units and how to schedule maintenance activities so that operating and maintenance costs can be minimized.

- Another principal activity of SOP is to undertake the study to identify whether or not system generation capacity is sufficient to meet the demand, taking in account outages of generation units. Because of the unpredictable nature of demand and generation availability over this long-time horizon, some accounting of the range of probable economic operation is necessary. This particularly applies to the likely mode of operation or time range of operation.

In the competitive market, profits must be realized in order to remain in business. To maximize profits, each power plant must conduct SOP that achieves the goal of efficiently operating its generating units in order to minimize the LCC while meeting the growing and periodically swinging customer demand. This leads to the following formulation for the operation planning problem:

Given a forecast of future customer demand and market price (spot price), establish a generation strategy that minimizes LCC over the planning period while meeting customer demand and that accounts for all relevant constraints such as technical, environmental, and contractual requirements [35].

With this formulation, it is assumed that in cases in which a power plant has insufficient resources to cover its customer demand, this can be done through purchases in the spot market leading to a balance between customer demand and the production and purchasing of power. This might involve a financial risk, but no liability in case of a national deficit.

Figure 9 illustrates the input variables, constraints, and output variables. Input variables

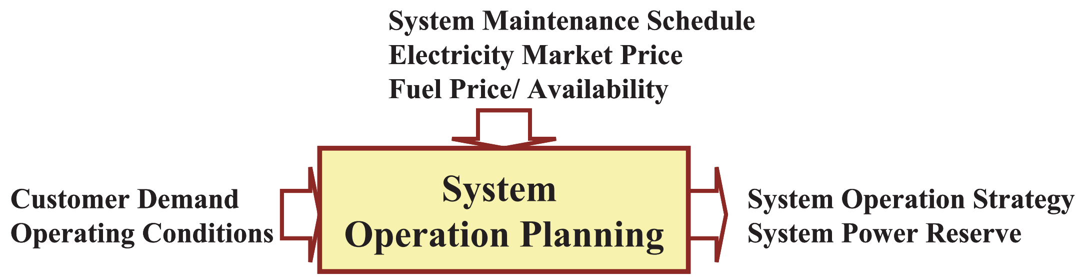


Figure 9: System Operation Planning Decision Action

include customer demand data, system operating conditions, fuel price and availability, and electricity prices. These data, which are different from the short-term forecast data in the OD and UC DAs, need to be forecasted on a long time horizon. Therefore input data involve a higher level of uncertainty. One input to this problem is the SMS identified by the SMS DA. Output data include the system operating strategies (SOS) and system power reserve.

2.1.5 System Capacity Expansion

The SCEP of power plants is an important, yet complex planning activity. As customer demand increases, the ability of a power plant to meet its customer demand decreases. Unlike most commodities, electricity cannot easily be stored, so it must be produced at the same instant it is consumed, and at the same time, it must be sufficient to accommodate the ever increasing demand of customers every second of the day and every day of the year. Recent blackouts in the western and eastern regions of the United States provide growing evidence that certain actions are urgently needed to ensure that power plants will continue to meet customer needs for reliable and affordable energy [85]. Much of the concern in this respect is due to the fact that the electricity infrastructure has made minute provisions to meet the changing needs of the economy. Therefore, to maintain an acceptable level of system reliability, the installed capacity of the power plant needs to grow to meet increasing customer demand by introducing new generation units.

SCEP is defined as the study of determining an investment plan for constructing generation units and interconnecting links; that is, its role is to determine where, when, and which generation units must be built and introduced into service, to guarantee an economical and reliable supply of the forecasted customer demand up to the horizon year. Again, two issues need to be considered simultaneously during the capacity expansion planning: economic issues and reliability issues [37].

Economic issues can be addressed by minimizing the expected sum of the investment and operation costs associated with each generation unit under uncertain conditions. The investment cost relates to the construction of generation units and interconnecting links, and is a function of the investment plan. The operating cost, which consists of maintenance costs, emission costs, fuel costs, and others, is mainly determined by fuel and maintenance costs of all the generation units. Efficient operation by managing these costs plays an important role during the life cycle of power plants.

The term "system reliability" has two aspects: system security, which measures the ability of a system to respond to disturbances arising within or outside of it; and system adequacy, which ensures that the system has sufficient capacity to satisfy customer demand. The reliability requirements ensure a balance between customer demand and production under various uncertain conditions [31]. Uncertainty stems from these sources [39]:

- Future operating conditions
  - Customer demand variations
  - Unit operating conditions
- Future social conditions
  - Construction time and constraints
  - Environmental constraints
- Future economic conditions
  - Fuel costs
  - Interest rates
  - Economic growth

The first step of SCEP can be carried out by identifying the power plant's initial information. Initial information includes the customer demand forecasting for the power plant, a set of new feasible generation units, the physical limits of each individual unit, and the cost of investment and operation. Because SCEP is a dynamic planning process and the decisions made earlier exert influences on the following stages, decision makers not only have to plan the system capacity expansion (SCE) for the whole horizon year but also analyze the system behavior for each planning stage as well [101]. In the first stage, the long-term SCEP is executed for the whole study time horizon, e.g., 20 years. In the second stage, mid-term planning is performed as the power plant approaches the target year, when much more precise information about the future are available. Mid-term planning involves analyzing and inspecting expansion plans previously identified in the long-term expansion planning. More accurate and detailed plans for the power plant become possible. At the final stage, as the power plant is closer to the target year, e.g., 5 years, short-term expansion planning takes place.

Traditionally, the deterministic approach has been used for SCEP with deterministic criteria. However, the probabilistic approach is more suitable for this long-term task that involves the need to represent in more detail some sources of uncertainty in future operating conditions, environmental conditions, and social conditions. The probabilistic approach is now widely used by power plants as an important method of incorporating uncertainty into operation and planning studies. Probabilistic-based criteria are also gradually replacing or supplementing deterministic ones. Studies [27] show that some sources of uncertainty have been more relevant to the DM process than others, but incorporating the various sources of uncertainty and accurately quantifying their impact, both in methodological and computational aspects, is an extremely complex task due to the following:

- In contrast with uncertainties in the operation condition, many of the uncertainties are strongly dependent on economics, politics, and social conditions. More general methodologies that analyze and quantify these uncertainties are required. Additionally, the way that the results are represented should more strongly emphasize discussion.

- The concept of a single capacity expansion plan is inadequate for performing system capacity expansion, necessitating the development of expansion strategies that account for possible evolutions of the power plant in different future scenarios and the dynamics of the DM process as uncertainties get resolved over time.

Figure 10 illustrates the input and output variables of this DA. Input variables include customer demand, operating conditions, and expansion cost data. Output variables include the number of new generation units, expansion costs and system capacity.

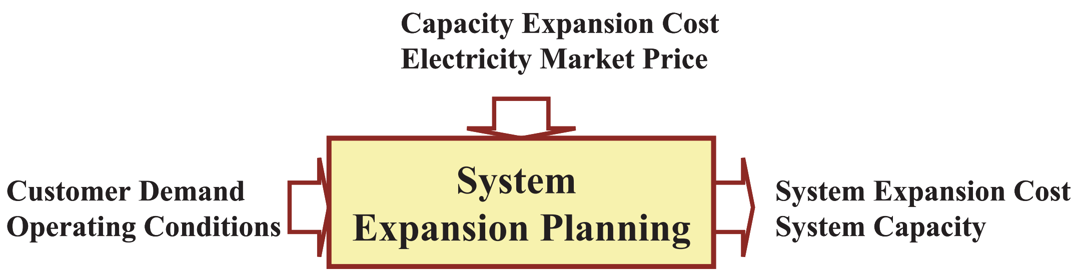


Figure 10: System Expansion Planning Decision Action

2.2 *Unit and System Maintenance Constraints*

To reduce the chances of trips and to minimize unscheduled maintenance of generation units, all power plants must schedule maintenance. Maintenance schedules not only determine maintenance costs but also affect system reliability and operating costs. Maintenance costs make up a significantly large percentage share of the total cost of power plants for the following reasons:

- The wide use of technologically sophisticated generation units requires higher levels of maintenance.
- The uncertainty caused by the use of sophisticated technologies in building generation units increases maintenance expenditures.
- Two major LCC components, capital costs and fuel costs, have decreased (see Figure 11). Capital costs are nearly 50% of what they were 10 years ago, and fuel costs

continue to decline because high-tech generation units work with higher efficiency and lower heat rates.

Today, maintenance expenditures can comprise between 15% and 20% of LCC [96].

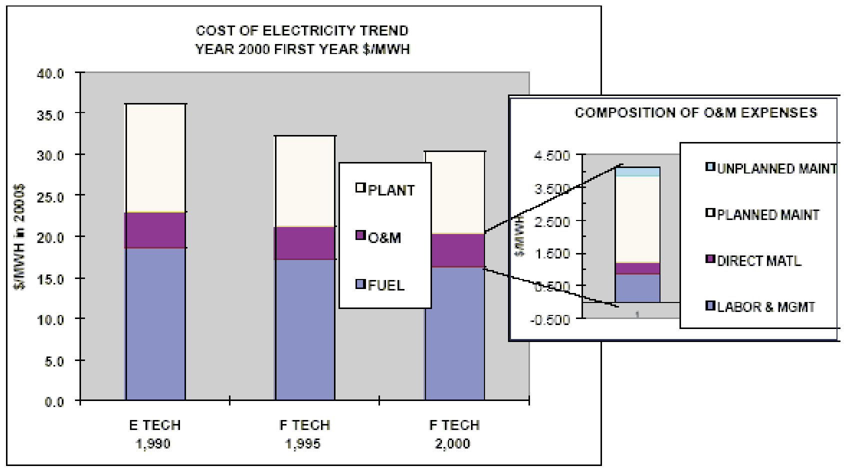


Figure 11: Trend of Maintenance Cost

Maintenance scheduling is a constrained optimization problem with a multitude of unit and system constraints that must be satisfied [113].

2.2.1 Unit Maintenance Constraints

Unit constraints are applied to the maintenance activities of each individual generation unit. These constraints may vary from one unit to another, depending on the properties of each generation unit and the type of maintenance required. Generally, these maintenance constraints can be categorized into four groups:

- **Maintenance window**: Defined as time slots when maintenance can be performed on generation units, the maintenance window for each unit specifies a time interval during which maintenance on that unit must take place and finish. The length of the maintenance window is determined by the type of maintenance activities that must be performed.

- **Crew constraints**: Crew constraints depend on human resources and their availability in the power plant. They specify the maximum number of units that can be in maintenance simultaneously without delay caused by a shortage of employees, as no two units can be maintained by the same crew simultaneously.

- **Resource constraints**: Because power plants have limited material resources and must keep inventory costs low, the resource constraints insure that no more than the available amount of resource for maintenance is committed. Otherwise, maintenance costs will skyrocket due to costs associated with delay, ordering, shipping, and materials.

- **Maintenance continuity**: Continuity of maintenance guarantees that the maintenance for each unit occupies the required time duration without interruption. The purpose of this constraint is to minimize both the unit's downtime and thus the downtime-related costs, especially during high-demand periods.

2.2.2 System Maintenance Constraints

System maintenance constraints are those constraints imposed on the generation units at the system level. Power plants must meet their customer demand reliably every second of the day and every day of the year, under normal, abnormal, and emergency conditions, including scheduled maintenance and unscheduled maintenance, guaranteed through load and reliability constraints.

- **Load constraints**: The demand requirements at any time are forecasted by the customer demand forecasting model. The units under maintenance contribute to lower system capacity. The most severe situation happens when the generation contingency is accompanied by a demand contingency. Appropriately adjusting the operating states of available generation units compensates for the loss of production due to maintenance, such as increasing the load level of some generation units from base load or part load to peaking load. However, this will hasten the wear of these units and thus increase the need for more frequent maintenance. Another way is to start up an off unit, if any exists, to remedy a loss of generation.

- **Reliability constraints**: Maintenance activities, on the one hand, play a very important role in the economical and reliable operation of power plants by extending the life of a generation unit and reducing the frequency of operation interruption. On the other hand, maintenance activities require careful planning and implementation. Increasing the load level of generation units in order to satisfy load constraints will certainly cause a high level of risk to the entire power plant, but system reliability requires that the system operate on an acceptable level of risk during the maintenance period.

The tradeoff between load and reliability constraints should be achieved so that power plants are to be operated more efficiently and responsively. The occurrence of unscheduled maintenance complicates this problem and makes it even harder to perform system planning due to its uncertainty.

2.3 Components Fired Factored Hours and Fired Factored Starts

Unit components wear down in different ways in different operating conditions, so regular inspections are important. Inspections provide direct benefits in reducing outages and increasing reliability, which in turn reduce unscheduled repair downtime.

2.3.1 Inspections

Three different inspections are usually performed: combustor inspection (CI), hot-gas-path inspection (HGPI), and major inspection (MI) [44]. Figure 12 shows the work scopes of these three inspections.

Compared to the other two inspections, the CI is a relatively short disassembly shutdown inspection. The scope of this inspection includes fuel nozzles, liners, transition pieces, crossfire tubes and retainers, spark plug assemblies, flame detectors, and combustor flow sleeves. Combustor liners, transition pieces, fuel nozzles and end caps are the focus of this type of inspection because they are usually the first to require replacement and repair. Proper inspection, maintenance, and repair of these items will contribute significantly to the longer life of the downstream parts, such as turbine nozzles and buckets.

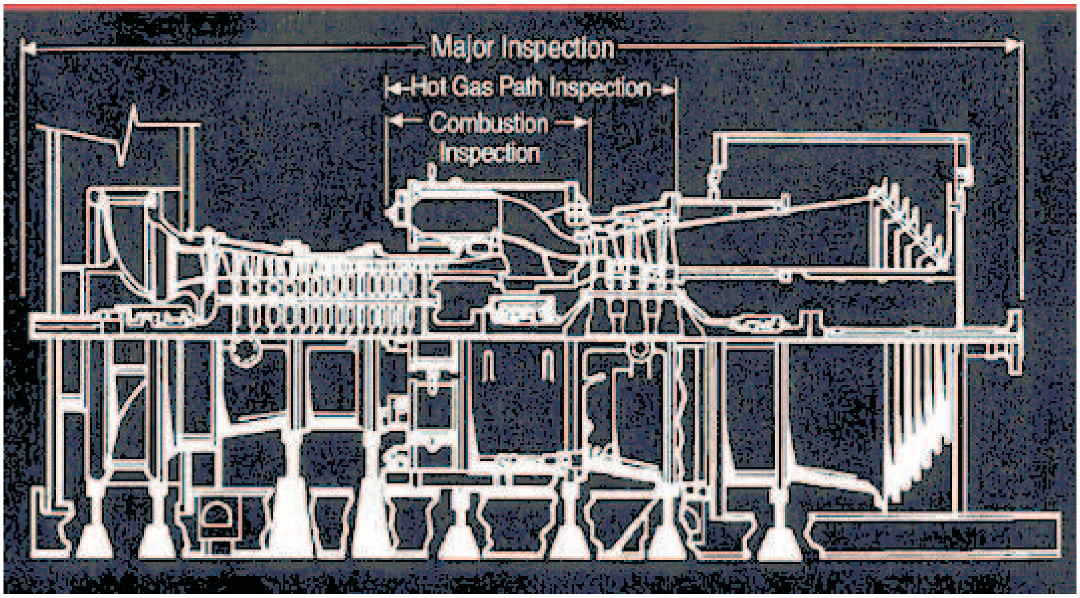


Figure 12: Different Inspections Work Scope

The gas discharged from combustion retains a very high temperature, exposing downstream parts such as turbine buckets, turbine nozzles, and stationary stator shrouds to continuous high temperatures. Thus, HGPI is needed for all these parts. The scope of this inspection includes all the components in the combustor and downstream from the combustor. CI is a part of the HGPI.

The scope of a MI includes all internal rotating and stationary components from the inlet through the exhaust section of a generation unit. It focuses on the inspection of all of the major flange-to-flange components of the generation units that are subjected to deterioration during the normal turbine operation. The MI includes previous CI and HGPI.

2.3.2 Duties

The maintenance requirements for each generation unit and each part of the generation unit are heavily dependent on the type of operation that the unit sees. For example, for a generation unit that constantly operates at peaking load, the dominant limiter is the thermal mechanical fatigue, but for a generation unit that operates continuously, the dominant life limiters are creep, oxidation, and corrosion. Generally, the operations of typical gas turbine application are categorized as peaking duty, cyclic duty, and continuous duty.

- Peaking duty is characterized by a relatively high starting frequency and a low number of operation hours per start. The seasonal variations of customer demand necessitates the peaking duty of some generation units. During high-demand periods, some units need to operate at peaking duty, while during low-demand period, they can be shut down. Thus, a high percentage of starts for peaking units are cold starts.

- Cyclic duty units start daily and typically operate twelve to sixteen hours per day. During weekends, the units are shut down due to lower customer demand. A large percentage of starts are warm starts due to the warm rotor condition. Cold starts occur when a start-up follows a two-day weekend shutdown, or a maintenance activity in which case the temperature of the units has become close to the ambient temperature.

- Continuous duty units undergo a low number of starts and a high number of operation hours per start. Most starts are cold because outages are generally maintenance driven. The maintenance requirements of continuous duty units are determined by the number of operation hours, not by starts.

Table 1 shows the different combinations of hot, warm, and hot starts, the operation hours per start for peaking duty, cyclic duty, and continuous duty, respectively.

Table 1: FA Gas Turbine Typical Operational Duties

Operation	Peaking	Cyclic	Coninuous
Hot Start (Down $<$ 4 Hr.)	3%	1%	10%
Warm 1 Start (Down 4 – 20 Hr.)	10%	82%	5 %
Warm 2 Start (Down 20 – 40 Hr.)	37%	13%	5%
Cold Start (Down $>$ 40 Hr.)	50%	4%	80%
Hours/Start	4	16	400
Hours/Year	600	4800	8200
Starts/Year	150	300	21
Percent Trips	3%	1%	20%
Number of Trips/Year	5	3	4

2.3.3 Fired Factor Hours/Starts

The gas turbine maintenance requirements for General Electric (GE) Power Systems are based on independent counts of starts and hours. Whichever criterion limit is first reached determines the maintenance interval. A graphical display of the GE approach is shown in Figure 13. In this figure, the recommended inspection interval is defined by a rectangle that is established by the starts and hours criteria. The recommended inspection should fall within the design life expectation. At the same time, it should be selected such that components are acceptable for continued use at the inspection point and will experience a low risk of failure during the subsequent operating interval. Replacement intervals are usually defined by a recommended number of inspection intervals and component specific (see Figure 14).

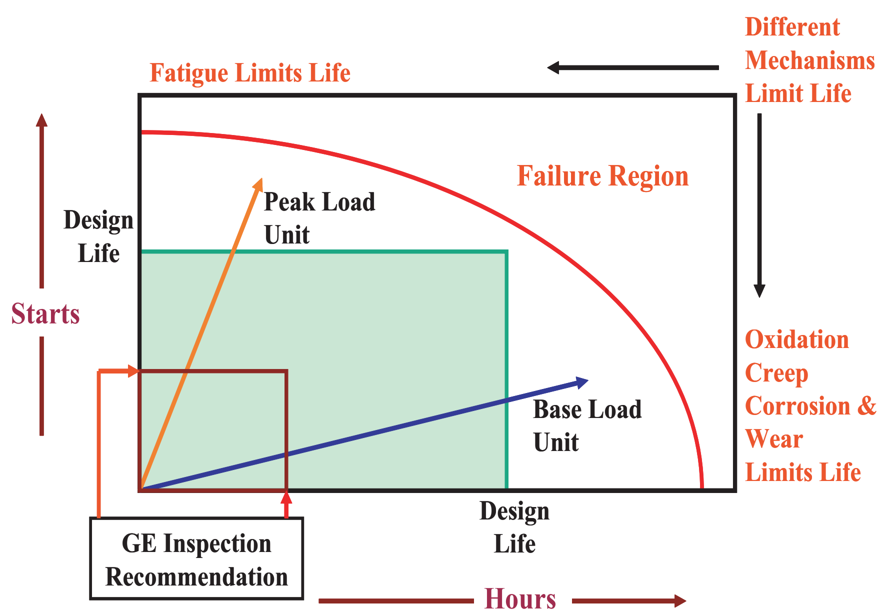


Figure 13: GE Bases Gas Turbine Maintenance Requirements on Independent Counts of Starts and Hours

By defining fired factored parameters, GE is better able to determine the appropriate maintenance intervals for their generation units. A parameter called "Fired Factored Hour" considers the impact of fuel type and quality, load setting, and steam or water injection. Another parameter, the "Fired Factored Start," considers the effect of the types of starts whether the generation unit is cold or hot and the rate at which the starts are taken. Ideal

PG7241 FA Parts

	Repair Interval	Replace Interval (hour)	Replace Interval (start)
Combustion Liners	CI	2 (CI) (1)(2)	5 (CI)(2)
Caps	CI	3 (CI) (2)	5 (CI) (2)
Transition Pieces	CI	3 (CI) (2)	5 (CI) (2)
Fuel Nozzles	CI	3 (CI) (2)	3 (CI) (2)
Crossfire Tubes	CI	2 (CI) (1)(2)	2 (CI) (2)
End Covers		4 (CI) (2)	3 (CI) (2)
Stage 1 Nozzles	HGPI	2 (HGPI) (3)	2 (HGPI) (3)
Stage 2 Nozzles	HGPI	2 (HGPI) (3)	2 (HGPI) (3)
Stage 3 Nozzles	HGPI	3 (HGPI)	3 (HGPI)
Stage 1 Shrouds	HGPI	2 (HGPI) (3)	2 (HGPI) (3)
Stage 2 Shrouds	HGPI	2 (HGPI) (3)	2 (HGPI) (3)
Stage 3 Shrouds	HGPI	3 (HGPI)	3 (HGPI)
Exhaust Diffuser	HGPI		
Stage 1 Bucket	HGPI	3 (HGPI)	2 (HGPI)
Stage 2 Bucket	HGPI	1 (HGPI) (4)	2 (HGPI) (5)
Stage 3 Bucket	HGPI	3 (HGPI) (6)	3 (HGPI)

CI = Combustor Inspection Interval

HGPI = Hot Gas Inspection Interval

(1) The goal is to increase this interval.

(2) Decision will be made based on fleet leader experience.

(3) The goal is to increase to 3 (HGPI). Decision will be made based on fleet leader experience.

(4) Interval can be increased to 2 (HGPI) by performing a repair operation. Consult your energy services representatives for details.

(5) Interval can be increased to 3 (HGPI) by performing a repair operation. Recoating at 1st HGPI may be required to achieve 3 (HGPI) replacement life.

(6) GE approved repair procedure at 2nd HGPI is required to meet 3 (HGPI) replacement life.

Figure 14: Estimated Repair and Replacement Cycles

operation is defined and used as a benchmark in measuring these influences. A generation unit in ideal operation operates on continuous duty with no water or steam injection.

2.3.3.1 Fired Factored Hours (FFH)

This parameter is an hours-based criterion, utilized to account for influences such as fuel type and quality, firing temperature setting, and the amount of steam or water injection, which reduce the maintenance intervals from the ideal case. In Figure 15, case 1 illustrates the impact of these non-ideal factors when they are involved in a unit's operating profile. The generation unit is operating for 8,000 hours, 160 starts per year. According to Table 1, this operating profile belongs to continuous duty. FFH are the determinant factor and three years is the maintenance interval for ideal operation. However, if the operation deviates from the ideal condition caused by either the firing temperature or steam/water injection, or fuel type, the maintenance criteria described by the rectangle for this operation decreases from the ideal case, e.g., the maintenance interval is reduced to two years.

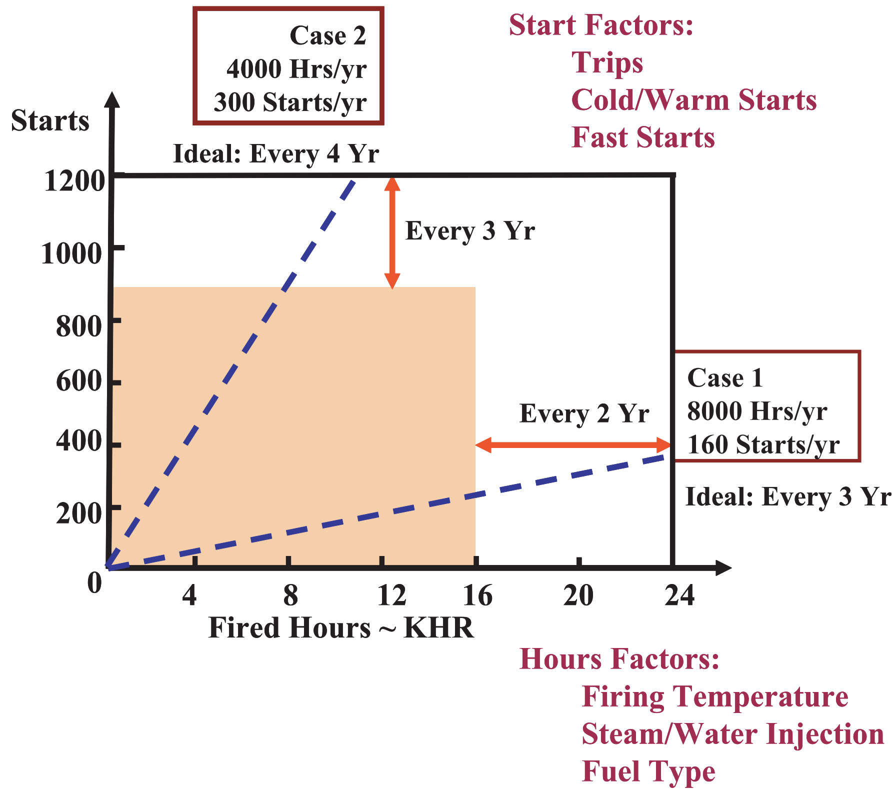


Figure 15: GE Maintenance Interval for Hot-Gas-Path Inspections

2.3.3.2 Fired Factored Starts (FFS)

This parameter is a starts-based criterion utilized to consider the impact of the start-up rate and the number of trips in the maintenance interval. FFS are determined for cold, warm, and hot starts over a defined time period by multiplying the appropriate cold, warm, and hot start operating factors by the number of cold, warm and hot starts, respectively. FFS for trips are also included. In both cases, these influences may act to reduce the maintenance intervals, also shown in Figure 15, case 2. The operating profile is 4,000 hours, 300 starts a year, which belongs to the cyclic duty. In this case, either FFH or FFS can be a determining factor for maintenance, depending on which one is reached first. After four years of ideal operation, FFS is first reached. The maintenance interval is reduced to three years for the operation that is different from the ideal one.

2.3.3.3 Maintenance Factors

A maintenance factor is defined as the ratio of a fired factored parameter to the actual value of that parameter. For the hours-based criterion, a maintenance factor is determined by dividing the FFH by the actual number of operating hours. Another maintenance factor based on starts is determined by dividing the FFS by the actual number of starts. Equations that determine application-specific hot-gas-path, combustor, and major inspections have been developed, see Appendix A. A maintenance factor is a number whose value is equal to one for the ideal operating condition or larger than one, in which cases, the inspection intervals are reduced from the ideal operating condition.

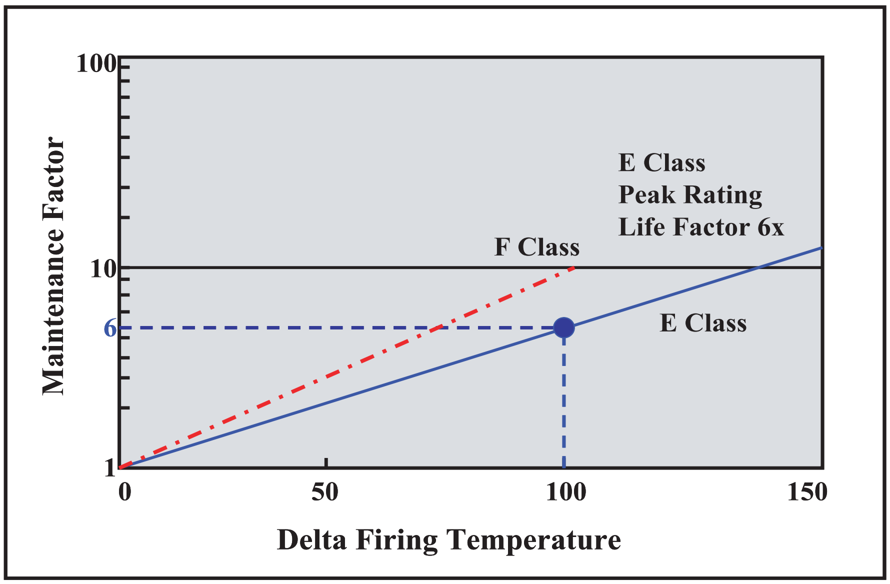


Figure 16: Bucket Life Firing Temperature Effect

For an MS7001EA turbine, each hour of operation at peak load firing temperature is the same as six hours of operation at base load from a bucket parts life standpoint and will result in a maintenance factor of six. Figure 16 defines the parts life effect corresponding to changes in the firing temperature. The significant operation at peak load will require more frequent maintenance and replacement of hot-gas-path components because of the higher operating temperatures. A higher firing temperature will reduce the lives of parts while a lower firing temperature will increase them. This provides an opportunity to balance the negative effects of peak load operation by periods of operation at part load.

2.4 Forecasting Variables

Any DM process requires that a sub-system provide information based on which decisions are made. This sub-system should include forecasting support among other capabilities. Moreover, it should be capable of integrating forecasting models into the DM process.

It is also true for the fleet management of power plants. The DM process heavily depends on the forecasting information on customer demand, fuel prices, and electricity prices. The forecasting information, together with the power plant operating data, which include the generation capacity, unit availability, and so forth, serve as input information for the DM process. As the input information is time dependent, it should be treated differently depending on the time horizon of the decision. The outputs from each DA act together in a complicated feedback and feedforward manner, causing the extensive inter-dependency of the decisions in the final operating characteristics of power plants. These operating characteristics can be measured by LCCs, profitability, reliability, or other gauges.

Various types of forecasting models have been employed. The choice of a method generally depends on the study horizon of the problem and the characteristics of the historical data in hand. In general, the time horizon is the most important factor since it determines which forecasting method proves the most effective. A second most important detail is the historical data, i.e., the number and the location of the sources of the data. In this section, the characteristics of the historical data of customer demand, natural gas prices, and electricity prices will be discussed in detail. Before doing so, the forecasting environment, the current electric market, will be reviewed.

2.4.1 Electric Market

Power plants are currently operating in a volatile market. This volatility is the result of the end of monopolies and the division of the electric business into several systems that manage the ares of generation, transmission, and distribution of electrical energy. This gives the electricity market a horizontal structure, unlike the traditional vertical structure. This process, known as deregulation, has lead to an open and competitive market that reacts in a similar manner to the stock market, but presents additional difficulties.

For one, the competitive market has brought about a high level uncertainty and risk. In the new framework of the electric market, major sources of uncertainty are market prices, customer demand forecasts, the availability of generation units, and other sources, many of which are dependent on each other or strongly correlated. These uncertainties adversely affect the underlying principle of the deregulated electric business, the efficient and full realization of existing generation sources, by introducing the risk of less secure power plants, unserved energy, and loss of opportunities [95]. Therefore, all sources need to be integrated in a unified framework in which risk and uncertainty are adequately addressed in a DM problem.

In this uncertain environment, the operations in the electric market must also adhere to all the physical rules involved in the process. One is to store significant amounts of electrical energy which indicates that the balance between production and demand should be maintained at all times [32]. In addition to the physical constraints, environmental behavior has to be introduced into the DM process. Hence, decisions must be made according to the expected behavior of electric markets and the physical rules of the power plants.

Therefore, deregulation has created a market that power plants have not yet adapted to. New regulations must replace the old ones before the implementation of the new market can be efficiently undertaken. Thus, while this is taking place, power plants must adopt new approaches that comply with the regulations.

2.4.2 Customer Demand Forecasting

Customer demand forecasting is defined as the forecasting of the amount of electricity that will be needed to supply a specific service area of customers. It can be categorized into short-term and long-term functions, depending on the horizon under consideration [115]. Short-term customer demand forecasting deals with hourly forecasting from one hour to a week ahead. Long-term forecasting usually covers forecasting horizons from one to ten years, and sometimes up to twenty years. In the electric business, long-term customer demand forecasts are primarily intended for capacity expansion, capital investment return studies, revenue analysis, fuel budgeting, and other issues. Unfortunately, accurately forecasting future events over long-term horizons poses great difficulty because of the innumerable uncertainties that characterize the underlying processes.

Customer demand forecasting information is one important input to many decisions made in power plants. Decision makers rely on forecasting to help improve the quality of their decisions regardless of wether it is operation or expansion planning. Accurate forecasts improve the efficiency of system operations by preventing unnecessary start-ups of generation units, scheduling suitable maintenance activities, and instantaneously delivering high quality electric energy to customers in a secure and economic manner whenever they need it. Accurate forecasts are also used by power plant management systems to establish capacity expansion plans for their systems. Therefore, the accuracy of the forecast strongly influences capital investment and is therefore imperative [107]. The quality of the forecasting separates wise investment decisions from poor ones. In the current electric market, customer demand forecasting must be improved.

Literature on various methods of generating accurate customer demand forecasts initially appeared over twenty years ago [60]. Since then, numerous methods that consider the time horizon of the forecast have been applied to the problem [100], [20], [68], and [56]. The most commonly used approach is the time series method. This method forecasts the future based solely on the assumption that the future will custom to that found in the historical data. However, this conventional technique fails to offer the level of accuracy and consistency that today's competitive market demands, as it does not allow the power plant to adapt to sudden changes with short durations in the business environment. In order to obtain high forecasting accuracy, more elaborate models must be developed.

Figure 17 shows the historical data of customer demand in a certain area for over 20 years. The data reflect the existence of seasonal patterns and a long-term trend of demand development, which is the critical information for forecasting. Furthermore, customer demand is affected by many other factors, so the forecasting process must also consider the impact of these accompanying conditions. Before reliable forecasting can be developed, several key issues must be addressed:

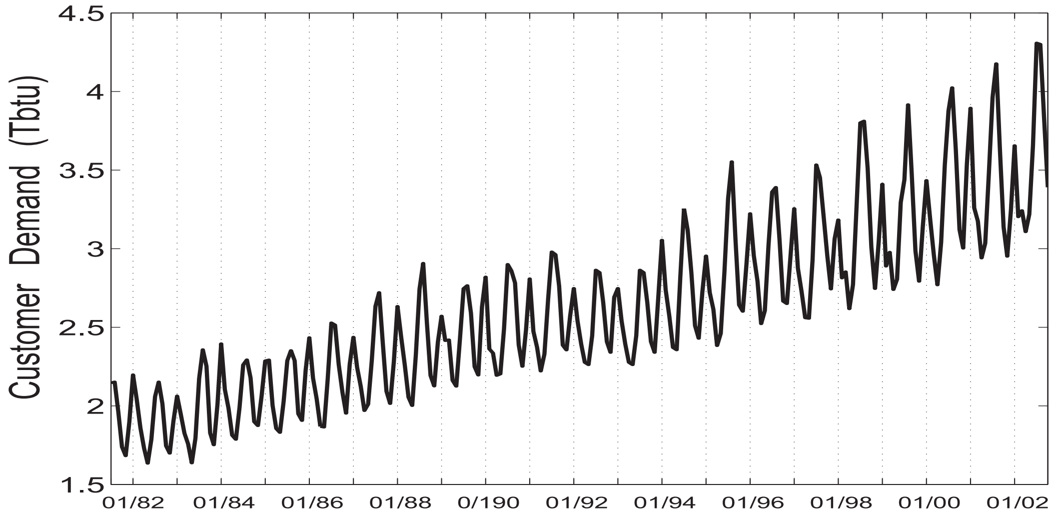


Figure 17: Historical Monthly Customer Demand

- For accompanying forecasting environment, the relevant variables with strong correlation to customer demand such as temperature, humidity, and wind must be identified and incorporated into the forecasting process.

- For historical data, a reliable data analysis and feature extraction technique that captures the dominant information related to patterns and profiles must be developed.

- The forecasting model must be able to extrapolate with a reasonable degree of accuracy when changes such as those involving socio-economic conditions or special events occur in the external business environment.

Finally, in such uncertain cases, decision makers or forecasters must utilize new data as they become available or make new assumptions for forecasting. Thus, the forecasting process has to be regularly re-evaluated and updated whenever new or relevant issues arise. This must also be done for electricity spot market price forecasting and fuel resource forecasting.

2.4.3 Electricity Spot-Market Price Forecasting

Electricity is not only a commodity but also an essential service, a key component in all other markets and business. Since customers will not tolerate less reliable service than what they are accustomed to from other commodities, the continued reliability of electric service must be examined thoroughly. Correspondingly, the price of electricity should be parsed into two components: the price of energy and that of reliability [15]. Because of these characteristics, the structure of electric markets differs from that of other commodity markets.

If the price of electricity is based solely on the price for electrical energy, the electricity market will not function properly. The emerging market for electricity has brought about an effort to price electricity on the basis of a charge for electrical energy, led by those who argue that the price would eventually stabilize at a marginal cost of electrical energy. The marginal cost is the supply curve that represents demand, which has been modeled as if there were no price elasticity, i.e., it is a vertical line. However, the price of electricity has not stabilized at the marginal cost of electrical energy. Figure 18 shows the historical electricity price.

From the point of view of demand, customers want electrical energy, but most customers also want reliability, i.e., they want electrical energy instantaneously upon demand, or they want a choice of level of reliability. Such reliability comes at an extra cost, passed on to customers who are paying for additional equipment that ensures that the power plant meets the demand instantaneously and that no demand remains unserved. A rule of thumb for power plants is that used and useful generation capacity should be 15% above the highest anticipated demand, justified because it ensures the reliable operation of power plants and a reliable supply of electricity. The problem in California occurred when the capacity was inadequate during a high demand period because customers were not charged for the cost of reliability. When the capacity of system did not meet demand, the price of reliability skyrocketed, exacerbating the problem.

From the point of view of supply, power plants must supply not only electrical energy but also reliability. That is, they must be able to balance supply and demand instantaneously. The cost associated with reliability, or the continuity of supply, differentiates electricity from other commodities. It includes the costs of spoiled or damaged products incurred by manufacturers, the costs of loss of business incurred by commercial business, the costs of

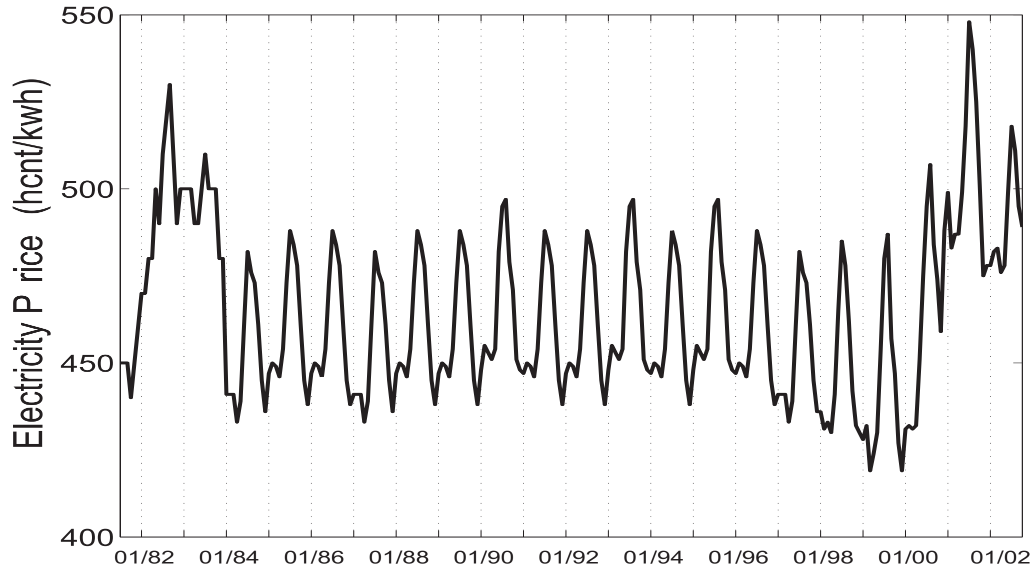


Figure 18: Historical Monthly Electricity Prices

health incurred by hospitals, the costs of traffic lights and other essential services incurred by city governments, and an inconvenience charge for all customers. The electric power industry has made a considerable effort in the past to quantify the cost or value of unserved energy, see [11] and [12].

Electricity prices play an important role in determining the value of generation units, wholesale contracts, and retail commitments, and therefore unambiguously set the current value of the power plant. Understanding this volatility facilitates the evaluation of different options conditional on those prices. Since usually few volatilities are known due to small amount of market information obtained in advance or lack of historical data, state of art electricity price forecasts have great business value. Such forecasts allow decision makers to take advantage of the tremendous profit opportunities associated with decisions to build, buy, or retire generation units, undertake long-term operation tasks, and lock in retail and wholesale customers with fixed prices for extended periods. Such forecasts also aid in recognizing the merits of these decisions and quantifying risks.

Numerous methods have been applied to electricity price forecasting, depending on the time horizon of the problem, see [90] and [65]. In developing an electricity price forecasting model, recognizing that the prices are inherently uncertain over time due to the uncertainty

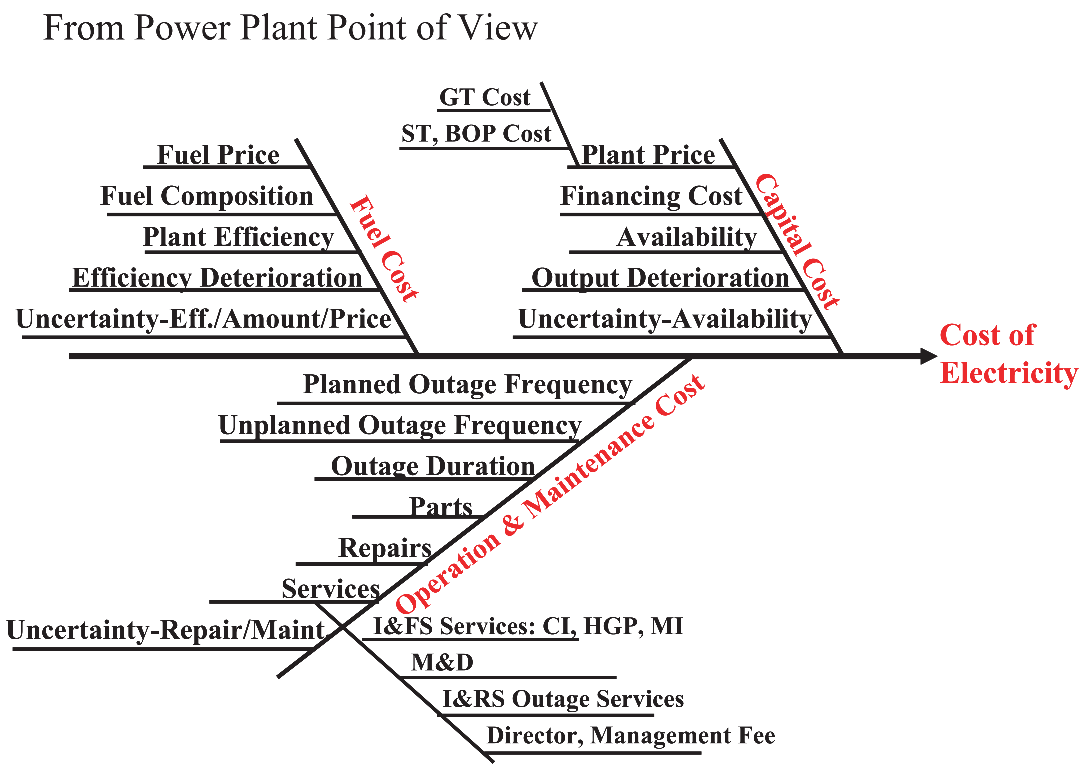


Figure 19: Factors Contributing to Cost of Electricity

in weather, generation units availability, fuel prices, and other related factors is critical. This uncertainty applies to all markets. Figure 19 shows the related factors that contribute to the volatility of electricity prices. Forecasting prices change as new information becomes available, which results in a forecast doomed to become obsolete. However, the "fact" of uncertainty (new information) does not obviate the usefulness of forecasts because new information can be utilized to complement the price forecasts rather than render them obsolete. Representing uncertainty in forecasting "qualifies" the forecasts so that the sensitivity of prices and their valuations to new information can be assessed. Thus, a model that accounts for the dynamics of the electricity prices is becoming increasingly relevant for power plants in the current electric market [15].

2.4.4 Fuel Requirement Forecasting

Fuel requirements and related forecasts are a key part of power plant planning. Fuel requirements include fuel availability, fuel prices, and fuel consumption. At most power plants, fuel accounts for 60 to 80 percent of operating costs, and for 20 to 40 percent of the total cost of electricity. Fuel expenditures are typically hundreds of millions of dollars a year [13]. Figure 20 illustrates the fraction of fuel cost to the total LCC for a typical combined-cycle power plant, see [116] and [49] for details.

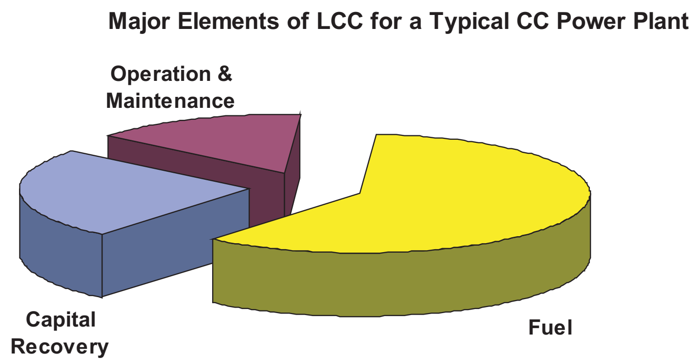


Figure 20: Fraction of Fuel Cost in the Total LCC of a Power Plant

Given the power plant cost structure, many cost elements are fixed and cannot easily be reduced significantly. Better fuel planning represents one of the few ways of reducing total life cycle costs of the power plant. It is also a critical input to many power plant decisions, such as fuel scheduling, contracting, ordering, and inventory planning. Thus, it represents one important way for power plants to maintain competitive in the current aggressive electric market. The major factors that need to be considered for fuel planning follow:

- Customer demand, representing the power plant production.
- Available system capacity, reflecting existing and planned system capacity.
- Unit availability, including unscheduled and scheduled outages.
- Unit dispatching, determining the loading level of each units.

As uncertainty in the power plant environment has increased in recent years due to fundamental changes in the business environment, each of these factors has become increasingly volatile. The sources of uncertainty can be grouped into two categories:

- Uncertainty in underlying long-term trends over time, such as the customer demand growth.

- Uncertainty in the short-term fluctuation around any given trend line, such as variations in annual customer demand due to economic or weather conditions.

Efforts to improve the ability to enhance fuel planning, especially fuel price forecasting, must address both kinds of uncertainty so that the forecast can reliably determine the expenditures resulting from the energy generation.

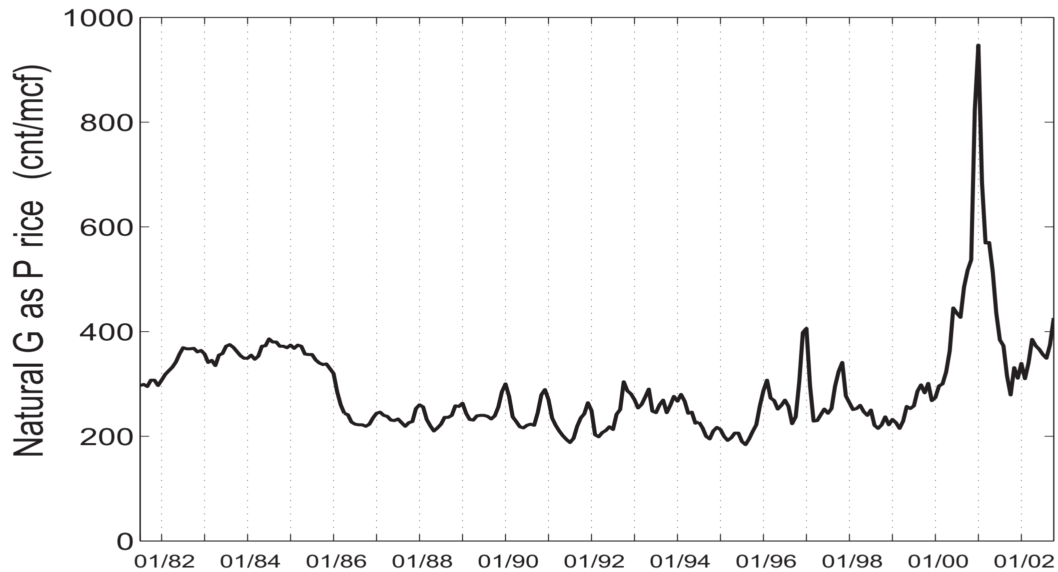


Figure 21: Historical Monthly Natural Gas Prices

The historical data of natural gas prices from July 1981 to October 2002 in Figure 21 show a big spike. Because traditional methods usually fail to capture such time localized phenomena, they have become inadequate. Without greater understanding of fuel requirement uncertainties and accurate forecasting, power plants face substantial risks such as that of overcontracting for fuel or committing to unprofitable bulk power transactions. Therefore, developing reliable forecasts has become extremely necessary in recent years.

2.5 Current Forecasting Methods

The above sections have explained why an accurate forecasting model is a very important information resource for decision makings in the electric power industry. The accuracy of such forecasts directly affects the validity of the decisions. Presently, forecasting methods in engineering are classified into the following four types:

- Qualitative Forecasting Methods
- Time Series Forecasting Methods
- Casual Forecasting Methods
- Simulation Forecasting Methods

2.5.1 Qualitative Forecasting Methods

Qualitative forecasting methods are primarily subjective methods that rely on human judgement and expert opinion. They are appropriate when little or no historical data are available, or when market intelligence is critical. In a newly emerging industry, such methods may be the only way to forecast several years into the future before sufficient historical data become available.

Qualitative forecasting methods are utilized primarily in two types of situations [110]. One method is used to forecast the time that a new process or product becomes widely adopted. For example, it is used in the forecasting of the point at which the application of a new scientific discovery becomes widespread, or in the prediction of the time horizon for the adoption of a new production process or development. A more specific example was the forecast of the time when laser technology would gain widespread industrial application. Quantitative forecasting methods would therefore be of interest to those organizations that have a widespread market for their product and the ability to exploit it. In this case, the timing of the development of products and marketing efforts that coincide with the demand for the products become the big concern.

The second situation that might require qualitative forecasting would be predicting what new developments and discoveries will be made in a specific area. For example, qualitative forecasting is used in the prediction of breakthroughs of medical research about some special disease, or in the prediction of the new technologies that will be developed in industry for the next several years that would help, say, perform SCEP for power plants.

Regardless of whether the forecast predicts the time at which some technologies will be adopted or the technologies and discoveries that will be made, quantitative methods based on the assumption that a pattern extracted from any available historical data is a good indicator of a future pattern cannot be used, as no such historical data are available. This gives rise to the need for qualitative forecasting methods.

Qualitative forecasting can be categorized into four methods:

- **Delphi Method**: Forecasting is developed by a panel of experts who anonymously answer a series of questions; responses are fed back to panel members who then may change their original answers. New group makes this process much more feasible. But this method is very time consuming and expensive.

- **Market Research**: Forecasting is done through questionnaires, or market tests, or surveys.

- **Product Life-Cycle Analogy**: Forecasting is derived from the life-cycles of similar products, services, or processes.

- **Expert Opinions**: Forecasts are based on the opinions provided by managers, sales force, or other knowledgeable persons.

Although recent years have witnessed considerable development in mathematical and statistical forecasting, it does not have to be quantitative. Many successful decisions are based on forecasts derived mostly from human judgement or expert experience and opinions. In these cases, mathematics and statistics work as tools that supplement sound business judgement.

The major difficulty in performing good qualitative forecasts usually arises when input is required from several executives working at different ranks in different departments of an organization. Low-level executives may have more access to critical knowledge of a product, but they may feel reluctant to speak up at a DM meeting if their ideas are in opposition to those of higher-level executives. In addition, people usually do not like to take extreme positions, but tend to moderate them so as to be closer to the mean, even if they foresee unusual patterns in the historical data. In either situation, qualitative forecasting may be difficult.

2.5.2 Time Series Forecasting Methods

A time series is a set of observations, each one recorded at a specific time [16]. Time series are often generated by monitoring industrial processes or tracking business metrics. Time series forecasting methods make use of these observed data to make a forecast. The use of time series is twofold:

- Obtain an understanding of the underlying forces and structure that produced the observed data. Data analysis should account for the internal structures of the given time series, such as the autocorrelation, trends, and seasonality. Usually, time series forecasting methods decompose the observations into both a systematic and a random component. The systematic component represents the expected value and consists of level (the current de-seasonalized value), trend (the rate of growth or decline between periods), and seasonality (the predictable seasonal fluctuations). The objective of time series forecasting methods is to filter out the random component and estimate the systematic component by using historical data. The random component is that part of the forecast that deviates from the systematic component, which cannot be forecasted with these methods.

- Fit a model and proceed to forecasting. It can be performed in two ways:

  - **Static forecasting methods** estimate the various parts of the systematic component once by utilizing the historical data, but they do not update these estimates on the value of the model parameters even though new information is observed. Static methods also assume that the initial estimates for the systematic component are correct and they treat all future forecast errors as a random component.

  - **Adaptive forecasting methods** update the estimates on the value of the model parameters of various parts of the systematic component after each new observation. Adaptive methods assume that a portion of the forecast errors are attributed to an incorrect estimation of the systematic component. Two popular adaptive forecasting methods are Holt's method and Holt-Winters' method.

Holt's model is appropriate when the time series has only level and trend but no seasonality in the systematic component. Holt-Winters' method is appropriate when the systematic component has not only level and trend but also seasonality as well.

The four steps in the adaptive forecasting framework are as follows [21]:

1. Initialization: Calculate initial estimates of the level, trend, and seasonal factors using the given historical data. This is done exactly as that in static forecasting methods.

2. Forecasting: Forecast for the period $t + 1$, given the estimates in period t, where $t = 0, 1, \ldots, n$.

3. Error Estimation: Record the actual value for period $t + 1$ and calculate the errors in the forecast for this period as the difference between the forecast and the actual value.

4. Modification: Modify the estimates of level, trend, and seasonal factors in the period $t+1$, given the forecasting errors. It is desirable that the modification be such that if the forecast is higher than the actual value, the estimates are revised downward; otherwise, the estimates are revised upward. The revised estimates in period $t + 1$ are then used to make a forecast for period $t + 2$ and steps 2, 3, and 4 are repeated until all historical data up to period n have been covered. The estimates at period n are then used to forecast the future value.

The advantage of time series forecasting methods is that they are easy to implement. They are applied in may fields such as economic and sales forecasting, budgetary analysis, inventory studies, workload projections, and utility studies. However, these methods assume that past history is a good indicator of the future, and the basic pattern does not vary significantly from one period to the next. Hence, these methods are most suitable when the business environment is stable. For more information on time series methods, see [40] and [99].

2.5.3 Casual Forecasting Methods

Time series forecasting methods do not explicitly identify the related factors that cause a particular movement in a time series over time. When experience and judgement are utilized for justifying changes in a time series caused by changes in one or more related factors, another avenue of forecasting is open, causal forecasting, which assumes that the forecasting variable is highly correlated with certain factors in the forecasting environment. Correlations between the forecasting variable and the related factors should be found first and then be utilized to perform forecasting. Therefore, the accuracy in forecasting the related factors determines the success of these methods.

Ideally, causal forecasting is used when the causal relationship is well-known and stable over time. Additionally, the causal (related) variables should be relatively easy to predict with high accuracy. For example, if a company that sells baby food wants to forecast sales for the next five years, the number of babies that will be born during each of the five years is a causal factor. A good forecast of this causal variable would be useful in forecasting the food demand. A highly accurate forecasting of the number of babies born in the United States should be possible by using Census Bureau data on the age distribution of the population, the average number of children born to each woman of child-bearing age, and other demographic variables [97].

The implementation of the causal forecasting should follow four steps [92]:

1. Regression: A mathematical equation relates a forecasting variable to one or more related factors that are believed to influence the forecasting variable.

2. Econometric Models: Interdependent regression equations describe activities such as economic activities in various fields.

3. Input-Output Models: The information flows describe the information from one field or sector to another. The outputs from another field or sector are required to predict the variables in this field.

4. Simulation modeling.

2.5.4 Simulation Methods

Simulation forecasting methods involve the use of analogs to model complex systems [106]. The analogs can take on several forms. A mechanical analog might use a wind tunnel to predict aircraft performance in real flight. A mathematical analog may use equations to predict the metric of interest, such as economic metrics. A metaphorical analog could involve using the growth of a bacteria colony to describe the growth of the human population. Game analogs are used for the interactions of players symbolic of social interactions.

Among these analogs, mathematical analogs are of particular importance and have been extremely successful in many forecasting applications, especially in the physical sciences. They are also used in the social sciences, but with lower accuracy, mainly due to the fact that social systems are usually extraordinarily complex. It is difficult to include all the related factors in a closed form model.

One of the most common mathematical analogs in quantifying societal growth is the S-curve. The model is based on the assumption of normal probability distribution. The process experiences exponential growth and reaches an upper asymptotic limit. Modis [61] has hypothesized that chaos-like states exist at the beginning and end of the S-curve. The disadvantage of utilizing the S-curve model is the difficulty in finding at any time a current location on the curve, or the proximity to the asymptotic limit.

Multivariate statistical techniques are often used in mathematical analogs in cases that involve relationships between two or more variables. Multiple regression analysis is the most commonly used technique, having become the primary forecasting tool in economics and social studies. It is different from trend extrapolation models, which only look at the history of the forecasted variable. Multiple regression models look at the relationship between the forecast variable and two or more related variables. It aims at understanding how a group of variables work together to affect another variable. In the multiple regression approach, as the correlations between the variables increase, the ability to predict any given variable decreases.

Another important simulation method is gaming analogs, in which players act according to a set of rules in an artificial environment or situation. Gaming has not yet been proven as a forecasting technique, but it does serve two important functions. First, by designing the game, variables of the system can be defined. Second, the relationships between the variables of the system can be studied.

CHAPTER III

APPROACH

A dynamic and adaptive modeling environment and methodology for the fleet management of power plants is going to be developed. The approach to developing this methodology is shown in Figure 22.

Figure 22 illustrates both the process of making decisions on the system level for a power plant that has a fleet of generation units and the relationships between each step in the process. The method begins with identifying the physical information of each generation unit and proceeds from the unit-level to the system-level characteristics. Maintenance scheduling, operational planning, and capacity expansion are the major long-term decisions that should be made on the system-level in order to operate a power plant both efficiently and responsively by fully utilizing its critical assets. Although the forecasting model is a sub-system, it is also the support system for the DM process. Because uncertainty is inherent in all systems, the uncertainty exploration must take place. Thus, this method will conclude by analyzing uncertainty so that decision makers are prepared for uncertainty in the future [77].

The focus of the current chapter is to address promising techniques and input needed to accomplish each step, to identify the interaction information, and to determine the output of each step that can be used in the DM process. A new forecasting method, which provide market information and support the DM process, is proposed. This chapter will yield a DM process for the fleet management of power plants that deals with "cross-scale" interactions, utilizes better market information, and evaluates the impact of pertinent external forces, it will also explore major uncertainty sources and provide comprehensive view and understanding of the developments of a power plant under different conditions.

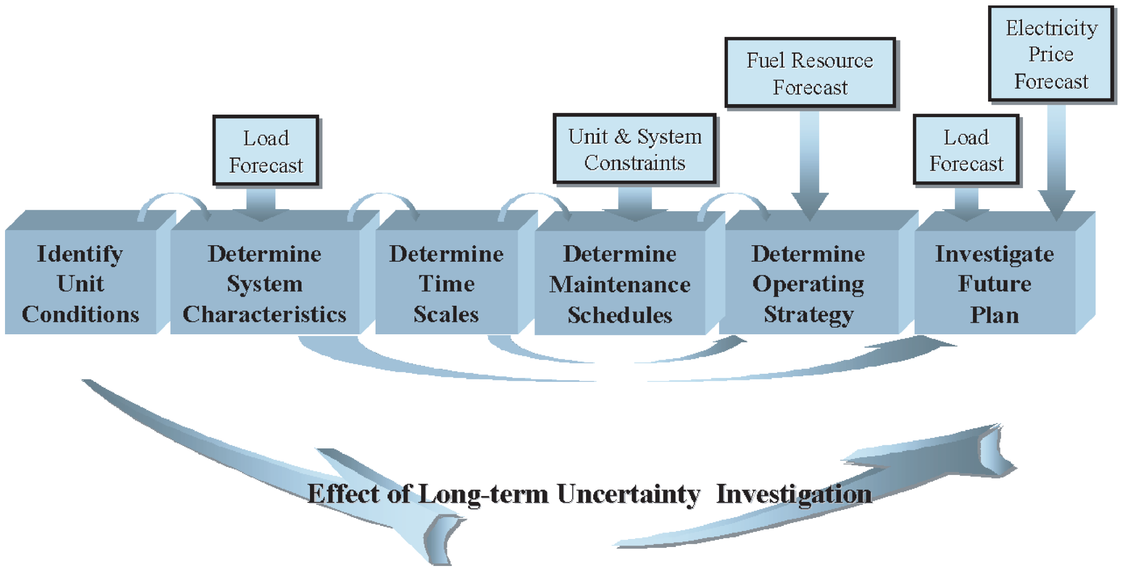


Figure 22: Flow Chart of the Modeling Methodology

3.1 Power Plant Fleet Management

3.1.1 Modeling and Simulation Environment

The modeling and simulation (M&S) environment is a valuable tool through which a better understanding of the system under study will be gained than that which could be achieved by merely solving an equation for the optimal value. An optimization formula may be easily applied to the system under steady state conditions. Electric power plants operate in an business environment that varies with time and exhibits characteristics of non-steady states. The M&S environment facilitates the study of such unsteady state systems by controlling certain conditions, accomplished through selecting variables that require changes and the ranges within which the changes vary. Different scenarios that describe the combinations of environmental and internal operation conditions can be generated. Future system conditions that assist and prepare the decision makers can be projected. M&S can also be utilized to "expand" time to "zoom in" on a certain event or "compress" time to gain a more comprehensive view.

From an economic viewpoint, the major advantage of using the M&S environment is that it is less expensive and involves less risk than actual experimentation. It would be much more expensive to change some aspects of the real world than to control variables in the

M&S environment. If some alternative failed or caused serious damage to the real system, it would be expensive or often impossible to restore the system to the original conditions so that another course of action might be taken.

The M&S environment is very important to electric power plants. With increased competition, power plants need to analyze different courses of actions searching for the better one. The M&S environment facilitates the evaluation of the impact of such courses of actions, which will assist decision makers to test alternatives before they are used in the real market. M&S can also be use to evaluate the influence of planned system changes such as capacity expansion or expected system changes such as growth in customer demand, changes in fuel prices, and so forth. In addition, it can assess the fluctuations in the business environment. M&S enables sensitivity analyses and evaluates the future performance of power plants under conditions of uncertainty.

3.1.2 Unit Operating Conditions

Unit operating conditions can be discretized into five conditions:

- Part load (oc_1)
- Base load (oc_2)
- Peak load (oc_3)
- Maintenance (oc_4)
- Off (oc_5)

Figure 23 illustrates the relationship between load setting and firing temperature. In the simple cycle mode, the turbine that maintains full open inlet guide vanes during a load reduction to 80% will experience a firing temperature reduction of over $200F$ at this output level [96]. The parts life under these various modes of operation can differ markedly. Significant operation at peak load will require more frequent maintenance and replacement of hot-gas-path components due to the high firing temperatures, which exacerbate creep, oxidation, and corrosion of the parts that are surrounded by high temperature gas, thus reducing parts lives. Lowering firing temperature will increases parts lives, providing an opportunity to balance the negative effects of peak load operation by periods of operation at part load or base load.

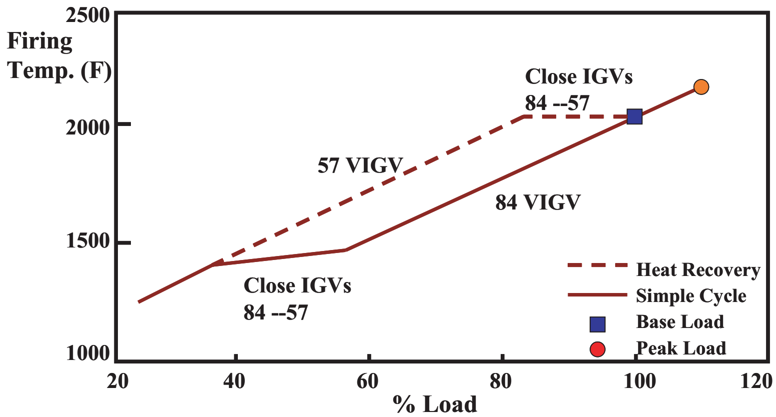


Figure 23: Load Setting and Firing Temperature Relationship for Simple Cycle Operation and Heat Recovery Operation

The need for maintenance is also dependent on the type of duty that a unit is operating at. In this study, continuous duty is assumed to be the duty type that the generation units will adopt without any specifications. As mentioned before, the maintenance requirements of continuous duty units will be determined by the number of operation hours, not by the number of starts. Thus, the "Maintenance" condition can be determined by the FFH of the components. Generation units that are either under "Maintenance" or "Off" are taken out of service and do not contribute to the generation of power.

At the beginning of a task, all the generation units are committable. That is, each generation unit is ready to be committed to produce power if customer demand is high and requires it to be committed. Thus, there are four operation conditions at the beginning of a task oc_s, where $s = 1, 2, 3, 4$. During the operation process, generation units can switch from one operating condition to another so that the total output can meet customer demand and that the total cost is minimized. Figure 24 shows the relationship between each operating condition. When one unit switches from "Maintenance" to any "Part load," "Base load," or "Peak load" condition, a start up cost is associated with the switching, $sc_{4,1}, sc_{4,2}$, and $sc_{4,3}$, respectively, where $sc_{i,j}$ represents the cost associated with switching

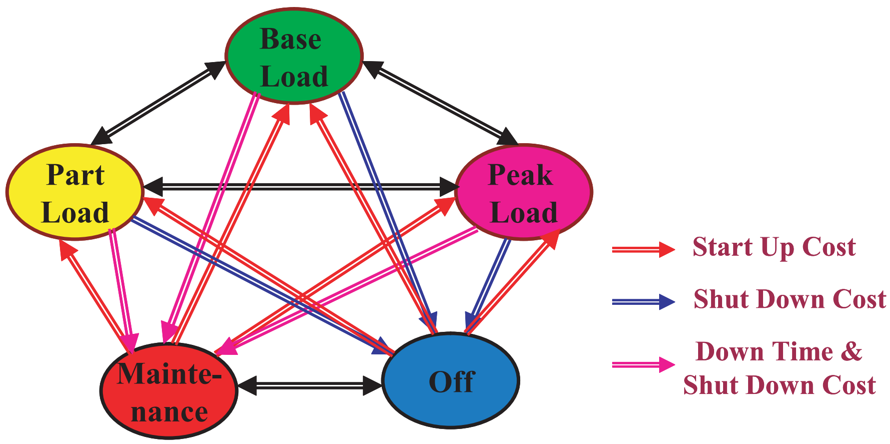


Figure 24: Operating Conditions

from operating condition oc_i to oc_j. The start-up cost is also associated with the condition switching from "Off" to "Part load," "Base load," or "Peak load," $sc_{5,1}, sc_{5,2}$, and $sc_{5,3}$, respectively. If any generation unit switches from generating power to either "Maintenance" or "Off," shut down costs occur, $sc_{s,4}$, or $sc_{s,5}$, where $s = 1, 2, 3$. For any generation unit that is committed but under "Maintenance" temporarily, maintenance costs and down time costs are incurred.

3.1.3 System Characteristics

The status of a power plant depends on the operating conditions of its generation units. At any point in time, the system status is determined by the operating conditions of each generation unit. A vector

$$SS_t = \{u_{1,s}, u_{2,s}, \ldots, u_{N,s}\}, \text{ where } s \in \{1, 2, \ldots, 5\}$$

can be used to describe the system status at time t, where $u_{n,s}$ denotes that unit n is in operating condition oc_s. A switch in the operating condition of any generation unit will change the status of the power plant. The power plant operating strategy is defined as the matrix

$$SOS = \{SS_t\}_{t=1}^{T},$$

which determines the operating conditions of each generation unit so that customer demand is satisfied at any time under any condition at a minimal total cost over the study horizon

T. SOS implies the system status SS_t must be adjusted from time to time. When the system output cannot meet customer demand, two situations need to be considered:

- The long-term situation. This situation is normally expected due to trends in ever increasing customer demand. The output of the power plant at time t, SO_t, might not be able to satisfy customer demand at time $t+1$. Even if it can, the output might not be optimal for achieving minimal cost. Therefore, SS_{t+1} needs to replace SS_t.

- The short-term situation. This situation can be either expected or unexpected. If at a point in time t^* that is within t and $t+1$, $t^* \in [t, t+1]$, a sudden increase in customer demand or maintenance (either scheduled or unscheduled) occurs, the system status SS_t must be changed. This case requires a finer time grid that allows quicker response to the demand contingency or the generation contingency.

The system output SO_t is determined by the number of units in operation in the power plant and the operating conditions for each individual unit.

$$SO_t = \sum_{n=1}^{N} O_{n,s}^t, \quad s \in \{1, 2, \ldots, 5\},$$

where $O_{n,s}^t$ is the output of generation unit n under operating condition oc_s at time t. System capacity is defined as the total output if each generating unit is operating at its base load condition. The system capacity is the output

$$SC = \sum_{n=1}^{N} O_{n,2}$$

under the system status $\{u_{1,2}, u_{2,2}, \ldots, u_{N,2}\}$. The system power reserve is defined as 20% of the system capacity. The system available capacity (SAC) is defined as the system capacity minus the system power reserve.

The customer demand forecasting model should provide the forecasted customer demand D_t over the planning time horizon. This information is needed not only to determine system status at any time step t, $\quad t = 1, 2, \ldots T$, but also to identify the economical operating time (EOT), which is determined to be the point in time such that

$$D_{EOT} = SAC.$$

That is, the time that the SAC meets the forecasted customer demand is defined as the economical operating time, and the period of time is defined as the economical operating period (EOP). Within the EOP, the power plant has no problem producing enough power to satisfy customer demand. The major concern is to find a SOS that minimizes the LCC, including maintenance and operating costs.

3.1.4 Identify Time Scales

The main difficulty with decision making in the power plant fleet management is to "zoom in" to "point events," such as maintenance activities, while considering the comparatively long-term operation process. Figure 1 shows that SOP and SMS have different time scales. If a decision is made with "point events" ignored, system status just needs to be updated by a certain time step t, depending on the rate at which customer demand increases and its seasonal variations. When "point events" that take place between time t and $t+1$ are taken into account, a finer time grid is needed to "zoom in" the time period between t and $t+1$ in order to respond to these events and update the system status with minimal delay.

A dual timescale system that replaces the single time scale traditionally used in the power plant fleet management should be utilized. Based on the frequency of each decision and the time frame during which that decision has an impact, a large time scale q is used for the determination of a SOS and a fine time scale w is used for the determination of SMS. System status is monitored for each fine time scale w. Power plants operate their generation units based on the time scale q. "Point events," such as maintenance activities or special events, act as a trigger that switches to the use of the fine time scale w. Therefore, during the period that "point events" occur, power plants operate their generation units based on the time scale w in order to quickly update the system status and minimize the costs associated with the "point events."

Customer demand, electricity prices, and natural gas prices, whose characteristics are illustrated in figures 17, 18, and 21, respectively, are the main inputs to the DM process. From these figures, all of these data series, particulary customer demand, clearly have seasonal variations. The determination of SOS should optimally capture the seasonality in customer demand. A quarter of a year has been selected as the time step for SOP. If no "point events" trigger the use of the fine time scale w, optimal system status SS_q will be updated for each q. During q, the power plant operates its generation units according to

$$SS_q = \{u_{1,s_q}, u_{2,s_q}, \ldots, u_{N,s_q}\}, \quad s_q \in \{1, 2, \ldots, 5\}.$$

"Optimal" indicates that this system status SS_q can satisfy customer demand during q at a minimal total cost. However, SS_q might be not optimal for the next period $q+1$ because of the variations caused by seasons or other factors. Thus, another system status SS_{q+1} that satisfies customer demand for this period should be selected. If customer demand changes so slowly that the SS_q can remain optimal, $SS_{q+1} = SS_q$.

A week w is selected as the time step for establishing the SMS, based on the fact that the maintenance window is usually in terms of weeks. Power plants need to operate their generation units differently because of generation contingencies. They must either increase the operating load level of other committed generation units or start up the "Off" generation units in order to compensate for the loss of generation due to one or several units under maintenance. The adjustment of the system status can be performed through w. When the system status is viewed through time step w, q is discretized into 13 segments. The system status at q can be expanded as a matrix with each row describing the system status at each week:

$$\mathbf{SS_q} = \begin{pmatrix} SS_{q_1} \\ SS_{q_2} \\ SS_{q_3} \\ SS_{q_4} \\ \vdots \\ SS_{q_{13}} \end{pmatrix}$$

with

$$SS_{q_i} = \{u_{1,s_{q_i}}, u_{2,s_{q_i}}, \ldots, u_{N,s_{q_i}}\}, \text{ where } s_{q_i} \in \{1, 2, \ldots, 5\}, \quad i = 1, 2, \ldots, 13.$$

If a generation contingency occurs at a certain w^* during a certain q^*, the power plant can respond to this contingency at $w^* + 1$, not $q^* + 1$. At the end of the contingency, the system returns to the original status which is optimized for normal operation conditions. For example, if a contingency occurs at $w^* = 3$, it will last for three weeks. For the system status during weeks 4, 5, and 6, $SS_{q_i^*}$ has switched to $\widetilde{SS_{q_i^*}}, \quad i = 4, 5, 6$, which are optimal and chosen to meet customer demand and minimize total cost in the maintenance window. After week 7, the system status returns to the one that is selected for this quarter. Thus, the power plant is able to operate under an optimal condition regardless of whether the system is under maintenance or not. This process is described in the following matrix:

$$\mathbf{SS_q} = \begin{pmatrix} SS_{q_1^*} \\ SS_{q_2^*} \\ SS_{q_3^*} \\ \widetilde{SS_{q_4^*}} \\ \widetilde{SS_{q_5^*}} \\ \widetilde{SS_{q_6^*}} \\ SS_{q_7^*} \\ SS_{q_8^*} \\ \vdots \\ SS_{q_{13}^*} \end{pmatrix}$$

One criterion is adopted to pick the optimal system status at each q and each w if there is a contingency. The selection criterion is defined as the ratio of the output of a unit at a certain operating condition for a given period of time to the FFH for that period of time, that is,

$$SM_{n,s} = \frac{O_{n,s}^t \times t}{FFH}, \text{ where } s \in \{1, 2, \ldots, 5\}, \quad n \in \{1, 2, \ldots, N\},$$

where $SM_{n,s}$ denotes the excellence value for generation unit n under operating condition oc_s. The higher the value of this parameter, the more efficiently the generation unit is operating at this condition than at the other conditions. It is desirable if all the generation units are operating at their most efficient operating conditions. Responsiveness requires that the power plant needs to satisfy customer demand. If these factors are considered, the most desirable system status can be determined.

The fine time scale is not only beneficial to dealing with scheduled maintenance but also to enabling power plants to react to unscheduled maintenance or unexpected increases in customer demand very quickly. Power plants can achieve these objectives through the weekly identification of the system status. The lead time that a power plant needs to react to "point events" is at most a small time step. Therefore, it provides a systematic mechanism of dealing with unscheduled occurrences.

Not only is the large time scale beneficial to capturing the seasonal characteristics of customer demand, electricity prices, and natural gas prices, but it is also capable of operating power plants more profitably. The system status is updated at the beginning of each quarter so that power plants can operate more efficiently. If this time scale includes more than one season, power plants try to function in the status that can satisfy the highest customer demand over the entire period. As we know, summer has the highest demand while spring is the season where the demand is relatively low. Hence, if the system status is selected based on demand in the summer, it will become inefficient during the fall and the winter, and especially during the spring. In contrast, if the system status is selected based on demand during the spring, then power plants will have very poor reliability of meeting customer demand during the other three seasons. These situations will definitely decrease the profitability of electric power plants.

3.1.5 Determine the System Operating Strategy

Take a power plant that has N generation units producing output over a planning horizon of Q, $(1 \leq q \leq Q)$, periods. For each unit, the unit capacity is denoted by C_n, $n \in \{1, 2, \ldots, N\}$. Under no circumstances can a unit's output exceed this limit C_n. A reserve capacity must be available in case of a unit breakdown or other unscheduled shutdowns. The forecasted customer demand in period q for the power plant as a whole is denoted by D_q, and the reserve capacity required by SR_q. The forecasted fuel cost is denoted by fc_q per unit output. Let mc_{n,w_q} be the maintenance cost of unit n if under maintenance at period w_q, which is the w^{th} week in the q^{th} quarter. Finally, let st_{n,w_q} be a state variable equal to one if unit n is being maintained in period w_q and otherwise zero.

$$st_{n,w_q} = \begin{cases} 1 & \text{if } n^{th} \text{ in maintenance in } w_q \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

In the EOP, power plant capacity is sufficient to meet the forecasted customer demand. The concern is how to operate the generation units so that the satisfaction of customer demands can be achieved with minimal total costs. Over the planning horizon $[1, Q]$, the objective of economical operation is to find a SOS that can supply the demand at the minimum total cost, which includes both maintenance and operating costs for the power plant. The maintenance cost is highly correlated with not only the SMSs, but also the unscheduled maintenance activities. The operating cost is mainly dependent on fuel prices and fuel consumption. The determination of SOS is classified as a cost-minimization problem that can be solved using an optimization-based technique.

The objective of minimizing the sum of the overall fuel and maintenance costs can be described as:

$$\text{Min: } \sum_{q=1}^{Q} \left\{ fc_q \left(\sum_{n=1}^{N} \sum_{w_q=1}^{13} O_{n,s}^{(w_q)} \right) + \sum_{n=1}^{N} \sum_{w_q=1}^{13} mc_{n,w_q} st_{n,w_q} \right\}, \quad s \in \{1, 2, \ldots, 5\}. \tag{2}$$

Various constraints need to be satisfied:

- The output of a generation unit must not exceed its capacity; the output is set to zero during maintenance:

$$0 \leq O_{n,s}^{(w_q)} \leq C_n (1 - st_{n,w_q}). \tag{3}$$

- The total output must equal the demand in each period:

$$\sum_{n=1}^{N} \sum_{w_q=1}^{13} O_{n,s}^{(w_q)} = D_q. \tag{4}$$

- The total capacity must not be less than the demand plus the required reserve:

$$\sum_{n=1}^{N} \sum_{w_q=1}^{13} (1 - st_{n,w_q}) C_n \geq (D_q + SR_q). \tag{5}$$

- Some units cannot be maintained simultaneously. If units n_i and n_j cannot undergo maintenance during the same period, this is represented by

$$st_{n_i,w_q} + st_{n_j,w_q} \leq 1, \quad i \neq j \in \{1, 2, \ldots, N\}. \tag{6}$$

- Unit n_i must be complete as a prerequisite for the maintenance for unit n_j to start. Let w^{n_i} denote the current maintenance starting period for unit n_i. This constraint can be represented as

$$w^{n_i} + mw_{n_i,p} \leq w^{n_j}, \text{ where } p \in \{1, 2, 3\} \text{ denotes the maintenance types.} \tag{7}$$

  In the case in which $w^{n_i} + mw_{n_i,p} \geq 13$, maintenance would continue at the start of the next quarter. Thus, a "wrap-around" plan would be generated.

- Once the maintenance of unit n starts, the generation units must be in a maintenance state for $mw_{n,p}$ contiguous periods for type p maintenance activity,

$$st_{n,w} = \begin{cases} 0 & \text{if } w = 1, 2, \ldots, w^n - 1 \\ 1 & \text{if } w = w^n, \ldots, w^n + mw_{n,p} \\ 0 & \text{if } w = w^n + mw_{n,p} + 1, \ldots, 13. \end{cases} \tag{8}$$

This optimization process is performed for the EOP. Based on the customer demand forecasted for the quarter, optimal system status SS_q is identified as a guide for how to operate the generation units. This quarterly update of the system status can better capture the seasonal characteristics and the growth rate of customer demand. Unit status is identified for each week based on the value of FFH for different parts of each generation unit to deal with scheduled maintenance. Any part that reaches the point required for inspection and maintenance will result in the generation unit being taken out of service. The maintenance window is determined by the type of maintenance activity p.

In cases of either generation or demand contingencies, system status needs to be updated to achieve a new "equilibrium" between production and demand. The system status can be adjusted in three different ways:

- The system output can still satisfy customer demand. System status remains the same.

- The system is not capable of producing customer demand.
  - Increase the operating conditions of some committed generation units to remedy the gap between the system output and customer demand.
  - Start up one or several "Off" generation units accompanied with adjusting the operating conditions of other committed generating units so that the system output can meet customer demand.

These choices will meet customer demand, but at different total costs. The choice depends on the total cost incurred. The one with minimal cost will be utilized to determine how to operate the generation units during the contingency periods. At the end of the contingencies, the system status returns to the original system status that is identified at the beginning of the current quarter.

3.1.6 Determine the System Maintenance Schedule

The maintenance activities can be categorized into two types: scheduled maintenance (preventive maintenance), and unscheduled maintenance (corrective maintenance).

Scheduled maintenance concerns with the scheduling of essential maintenance over a fixed planning horizon for a number of generation units while minimizing maintenance costs and providing enough generation to meet the anticipated demand. The unit status is determined by the previous operation of the unit. FFH accumulate from the time the parts go into operation. The unit status is determined for each small time scale based on the accumulative FFH and its the limit value $FFH_{L,p}$. The recommended scheduled maintenance should be performed for unit n on type p maintenance from the $w_q^n + 1$ week for $mw_{n,p}$ weeks if

$$FFH_{L,p} - \triangle FFH_p^n \leq FFH_{w_q,p}^n \leq FFH_{L,p},$$

where $\triangle FFH_p^n$ is the incremental value of FFH for type p inspection during one small time scale. The calculations of FFH for different components (see Appendix A) determine the recommended maintenance schedules for them. Modification of the recommended maintenance schedule must take into account the following conditions so that it can be carried out practically:

- Due to the limited resources of a power plant and the need to satisfy customer demand at any time, the maximum number of units that can be under scheduled maintenance simultaneously must be limited.

$$\sum_{n=1}^{N} st_{n,w^*} \leq M_{max}.$$

  At any week w^*, the number of generation units that can be under maintenance should be less than M_{max}, which is determined by the system resources.

- If two units are operating in a similar way, the value of FFH may be very close to each other. At some point in the operation process, these two units may require scheduled maintenance at the same time. The value of FFH alone cannot separate these two maintenance activities. The knowledge of the incremental value of FFH, $\triangle FFH$, for each small time period is used to switch one unit to a maintenance period ahead of the recommended scheduled maintenance so that the maintenance windows of these two units are seperable. If

$$FFH_{w_q,p}^{n_i} \geq FFH_{w_q,p}^{n_j},$$

  and unit n_i is in maintenance from $w_q + 1$ for $mw_{n,p}$ weeks if

$$FFH_{L,p} - (mw_{n,p} + 1) * \triangle FFH_p^{n_i} \leq FFH_{w_q,p}^{n_i} \leq FFH_{L,p} - mw_{n,p} * \triangle FFH_p^{n_i}.$$

  If this is done, the maintenance for unit n_i is usually brought forward by one maintenance window as compared with the recommended maintenance schedule. In this case, unit n_j can be maintained according to the recommended maintenance schedule.

The scheduled maintenance activity is a function of the maintenance window, the labor fees, material costs, downtime costs, and start up costs. During w, the maintenance cost for unit n can be determined by the following function if it is under maintenance:

$$mc_{n,w_q} = f(mw_{n,p}, LF_{n,p}, RC_{n,p}, DTC_{n,p}, STC_{n,p}),$$

where mc denotes the maintenance cost, LF stands for the labor fees, RC stands for material costs, DTC denotes downtime costs, and STC denotes startup costs. If a generation unit is suddenly under unscheduled maintenance, the small time scale mechanism enables the power plant to respond to this unscheduled event very quickly, e.g., generating unit n^* encounters an unscheduled maintenance from week w^{n*} in quarter q^* with maintenance window mw_{n^*,p^*}.

$$st^u_{n^*,w} = \begin{cases} 0 & \text{if } w = 1, 2, \ldots, w^{n*} - 1 \\ 1 & \text{if } w = w^{n*}, \ldots, w^{n*} + mw_{n^*,p^*} \\ 0 & \text{if } w = w^{n*} + mw_{n^*,p^*} + 1, \ldots, 13 \end{cases} \tag{9}$$

Several scenarios may occur:

1. If at the time the unscheduled maintenance occurs, there is no scheduled maintenance, this unscheduled maintenance activity can be treated as a scheduled maintenance from both generation resource and maintenance resource points of view. This can be precisely described by

   $$st_{n,w} = 0 \text{ when } w = w^{n*}, \ldots, w^{n*} + mw_{n^*,p^*}, \text{ where } n \in \{1, 2, \ldots, N\} \backslash n^*. \tag{10}$$

   Maintenance cost can be described as

   $$mc_{n^*,w^*} = f(mw_{n^*,p^*}, LF_{n^*,p^*}, RC_{n^*,p^*}, DTC_{n^*,p^*}, STC_{n^*,p^*}).$$

2. If at the time the unscheduled maintenance occurs, a future scheduled maintenance might be performed at the same time for the same unit. This situation can be treated as one maintenance activity

   $$st^{(s)}_{n^*,w} = 1 \text{ when } w = w^{n*}, \ldots, w^{n*} + mw_{n^*,p^*}. \tag{11}$$

   Maintenance cost can be described as

   $$mc_{n^*,w^*} = f(mw_{n^*,p^*}, LF_{n^*,p^*}, RC_{n^*,p^*}, DTC_{n^*,p^*}, STC_{n^*,p^*}).$$

3. The most challenging condition is one in which a scheduled maintenance is in process but for different reasons, causing a conflict in resource allocations.

   $$st_{n,w} = 1 \text{ when } w = w^n, \ldots, w^n + mw_{n,p}, \tag{12}$$

with

$$[w^n, w^n + mw_{n,p}] \cap [w^{n*}, w^{n*} + mw_{n^*,p^*}] \neq \emptyset,$$

and n, p satisfy the following conditions:

$$\begin{cases} n = n^*, p \neq p^*, \text{ or} \\ n \neq n^*, p = p^*, \text{ or} \\ n \neq n^*, p \neq p^*. \end{cases} \tag{13}$$

The maintenance activity will be restricted by the unit and system maintenance constraints. With regard to crew resources, the crew's response causes a time delay. Time delay also results when some parts needed in emergency are not stocked. However, such situations can be simplified. When a conflict between scheduled maintenance and unscheduled maintenance occurs, the electric power plant can perform both at the same time, while incurring overtime, shipping, order, and material costs. Despite the added costs, the power plant will adopt this course of action because it can shorten the time that two units are under maintenance, especially during high-demand periods.

3.1.7 Investigate the System Capacity Expansion Plan

The planning of the expansion of generation capacity is a complex process that involves the identification of future scenarios in terms of customer requirements, technical innovations, costs of capital and operations, economic and regulatory environments, and their interactions. Such planning becomes a major concern when the time of interest is beyond the EOP, and power plant capacity is less than the forecasted customer demand. Based on the information provided by the customer demand forecast model, this shortage of power can be considered either short term or long term.

A short-term shortage of power indicates after a temporary high-demand period, the power plant is still capable of producing enough power to satisfy customer demand. Then the power plant will chose to temporarily purchase electricity from other power plants to meet customer demand and to avoid even a higher penalty. The electricity spot market prices forecasting model will provide the information on the prices of electricity ec_q. In this case, the total cost of the system is the minimum maintenance and operating costs of the power plant operating at system capacity plus the cost of purchasing a certain amount of additional electricity that cannot be provided by the power plant.

$$\text{Min: } \sum_{q=1}^{Q} \left\{ fc_q \left(\sum_{n=1}^{N} \sum_{w_q}^{13} O_{n,s}^{(w_q)} \right) + ec_q \left(D_q - \sum_{n=1}^{N} \sum_{w_q}^{13} O_{n,s}^{(w_q)} \right) + \sum_{n=1}^{N} \sum_{w_q=1}^{13} mc_{n,w_q} st_{n,w_q} \right\} \tag{14}$$

A long-term shortage of power occurs when the forecasting information indicates that customer demand will continue to increase for a long period of time. In this case, decision makers at the power plant should consider expanding the generation capacity, the objective being to determine the number of units that should be added to the existing electric power plant to realize the expected EOP. The number of generation units added is determined based on the EOP and accompanying capacity the power plant expects after expansion. Here the total cost includes the maintenance cost, the operating cost, and the depreciation of investment (capital) cost cc_q.

$$\text{Min: } \sum_{q=Q+1}^{Q^E} \left\{ \left(cc_q + fc_q \sum_{n=1}^{N^E} \sum_{w_q=1}^{13} O_{n,s}^{(w_q)} \right) + \sum_{n=1}^{N^E} \sum_{w=1}^{13} mc_{n,w_q} s_{n,w_q} \right\}. \tag{15}$$

The number of generation units that needs to be introduced into the power plant $N^E - N$ is determined by:

$$\sum_{n=1}^{N^E} \sum_{w=1}^{13} (1 - s_{n,w}) C_n \geq (D_{Q^E} + SR_{Q^E}), \tag{16}$$

where Q^E is the expected EOP for the expansion power plant and N^E is the total number of generation units the power plant has after capacity expansion.

For the power plant with new generation units, the determination of SOS and SMS can be carried out in the same way as it was by the old power plant. The constraints from Equations (3) to (8) are applied to the expansion problem. Hence, the SCE can be directly integrated into the existing power plant plan.

3.2 Analysis of Electric Market Dynamics

The ability to forecast, in either an implicit or explicit form, is crucial for an information-decision-action system operating in uncertain environment. Strategic and operational decision making may depend heavily on the future conditions of the electric market provided by forecasting models. Therefore, an understanding of the behavior of the electric market is a critical task of decision makers. A large amount of data, such as customer demand, electricity prices, and fuel prices, which describe the behavior and properties of the electric market is available to decision makers, so they face the challenge of correctly interpreting such data and extracting critical information from them.

Historical data obtained in the electric market, a time-domain data series in the raw format, measure a function of time. However, this format is not always the best representation of the data for most data processing-related applications. Conventional approaches to analysis usually provide the best results for a stationary time series. When a series is non-stationary, as is the case for most time series in the electric market, a mechanism that reveals aspects of the data series that conventional techniques usually miss must be identified. An effective mechanism for such a task is a high-frequency filtering, seasonality identification, and trend analysis method, which enables the analysis of large volumes of historical data existing in the electric market, which in turn will render critical information that is not readily available in the raw format [73]. An efficient way is to utilize multi-resolution decomposition techniques such as the wavelet transform, which can produce a good local representation of the data in both the time and the frequency domains. In contrast to the Fourier basis, wavelets can be supported on an arbitrarily small closed interval. Thus, the wavelet transform is a very powerful tool for dealing with transient phenomena typical in the electric market. Combining wavelet transform in the historical data analysis and hybrid forecasting scheme can provide high accuracy forecasting results for the electric business.

This section will briefly introduce Fourier transform and then discuss the multi-resolution analysis and wavelet transform, which is a better way to represent time series.

3.2.1 Fourier Transform

In the 19^{th} century, the French mathematician, J. Fourier, showed that any periodic function can be expressed as an infinite sum of periodic complex exponential functions. This idea was first generalized to non-periodic functions and then to periodic or non-periodic discrete time series. These generalizations and the development of the Fast Fourier Transform (FFT) in 1965 made it become very popular and suitable for computer calculations [73]. Even now, the Fourier transform is probably by far the most popular among a number of transforms.

Fourier transform (FT) reveals the frequency content of a time series by decomposing it into complex exponential functions of different frequencies. In many cases, the most distinguished information is hidden in the frequency spectrum of a time series that shows what frequencies exist in it. The way it does this is defined by the following two equations:

$$X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt \tag{17}$$

$$x(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) e^{i\omega t} d\omega, \tag{18}$$

where t stands for time, ω stands for frequency, x denotes the data series in the time domain, and X denotes the data series in the frequency domain. Equation (17) is called the FT of $x(t)$, and Equation (18) is the inverse FT of $X(\omega)$.

The information provided by the integral Equation (17) corresponds to all time instances. The result of this integration will be affected equally regardless of where in time the component with frequency "ω" appears. Therefore, the FT tells if or how much a certain frequency component exists in a data series, but it does not provide information about when in time the frequency component exists. This is why FT is applicable to stationary data series whose frequency components do not change in time, but it is not suitable for non-stationary data series. The existence of a non-stationary data series, such as the historical data in the electric market, has necessitated the development of other transforms that can provide time-frequency representation (TFR).

3.2.1.1 Short-Time Fourier Transform

The Short-Time Fourier transform (STFT), or Windowed Fourier Transform, is a revised version of the FT. The data series is divided into small enough segments so that during which it can be assumed to be stationary. A window function "W" whose width must be equal to the width of the segments where its stationarity is valid is chosen.

$$STFT_X^{(W)}(t,\omega) = \int_t [x(t)W^*(t-\tau)]e^{-i\omega t}dt, \tag{19}$$

where $x(t)$ is the data, $W(t)$ is the window function, and $*$ is the complex conjugate. Thus, the STFT is a function of both time and frequency, which provides a TFR of the data series. STFT has the problem of resolution and selection of window function. The root of this problem is the well known Heisenberg Uncertainty Principle, which states that the information about the time and frequency cannot be obtained exactly simultaneously. One cannot know what spectral components exist at what instances of times. What one can know are the time intervals in which certain bands of frequencies exist. This is a resolution problem.

FT does not have a resolution problem in both the frequency and time domains, because what frequencies exist in the frequency domain and what the value of the data at every instance of time are precisely known. Conversely, the time resolution in the frequency domain and the frequency resolution in the time domain are zero, since there is no information about them. Because the window function used in FT lasts at all times, frequency resolution in the frequency domain is known perfectly.

In STFT, the window function is of finite length, covering only a portion of the data series. This causes the frequency resolution to get poorer. The exact frequency components that exist in the data series cannot be precisely known. Only a band of frequencies can be known. However, in order to apply stationarity, a short enough window is a must. The narrower (more compactly supported) the window, the better the time resolution; and the better the assumption of stationarity, but poorer the frequency resolution, and vice versa. For more on STFT, see [48] and [59].

Wavelet transforms, which provide the TFR simultaneously, were developed as an alternative to the STFT. Multi-resolution analysis (MRA) and wavelet transforms (WT) are addressed next.

3.2.2 Multi-Resolution Analysis

The time-frequency resolution problems are the results of physical phenomena and exist regardless of the transform used, but a data series can be analyzed by using an alternative approach: MRA. MRA analyzes the data series at different frequencies with different resolutions which differs from STFT, which resolves equally every spectral component. MRA gives good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies. Therefore, MRA is appropriate, particularly when the data have high-frequency components for short durations and low frequency components for long durations. Fortunately, the data series encountered in practical applications are often of this type, such as historical data in the electric market.

In MRA, a data series or a function can be viewed as composed of a smooth background and details on top of it. The distinction between them is determined by the resolution or by the scale below which the details of a data series can not be discerned. At a given resolution, a data series is approximated by ignoring all the details below that scale. Finer details are added to the coarser description by progressively increasing the resolution. This provides a successive approach to approximating the data series and finally recovering the data when the resolution goes to infinity.

In the space of square-integrable functions $\mathbb{L}^2(\mathbb{R})$, a sequence of resolutions labeled by the integers is defined such that all details of the data series on scales smaller than 2^{-j} are suppressed at resolution j. MRA decomposes the function space into a sequence of subspace V_j, which is the subspace of functions that contains data information down to scale 2^{-j}. The subspace V_j is contained in all the higher subspaces $V_j \subset V_{j+1}$ for all j; that is, the information at resolution level j is necessarily included in the information at a higher resolution, the first requirement for MRA. Let W_j be the detail space at resolution level j and orthogonal to V_j. The relationship between V_{j+1} and V_j can be expressed by the following equation:

$$V_{j+1} = V_j \oplus W_j. \tag{20}$$

This decomposition of the V_{j+1} space can be continued as

$$V_{j+1} = W_j \oplus V_j = W_j \oplus W_{j-1} \oplus V_{j-1} = ... = W_j \oplus W_{j-1} \oplus W_{j-2} \oplus ... \oplus W_{j-J} \oplus V_{j-J}. \tag{21}$$

Then the subspace V_j at resolution j can be expressed as a sum of subspaces that are mutually orthogonal, since $W_j \perp V_j$, W_j is orthogonal to any subspaces of V_j.

The second requirement for MRA is that all square integrable functions be included at the finest resolution and only the zero function at the coarsest level. As the resolution gets coarser and coarser, more and more details are removed from the data. At the limit $j \to -\infty$, only a constant survives. Since this constant must be square integrable, it can be only a zero function. On the other hand, if the resolution increases, more and more details are added. At the limit $j \to \infty$, the entire space $\mathbb{L}^2$ should be recovered; that is, $\lim_{j\to\infty} V_j = \mathbb{L}^2(\mathbb{R})$.

The third requirement for MRA is scale or dilation invariance. Subspaces V_j are scaled versions of the central space V_0. If $x(t) \in V_j$ contains no details at scales smaller than $1/2^j$, $x(2t)$ is a function obtained by squeezing $x(t)$ by a factor of 2, which contains no details at scales smaller than $1/2^{j+1}$. Therefore, $x(2t) \in V_{j+1}$.

The fourth requirement for MRA is translation or shift invariance. If $x(t) \in V_0$, so do its translates $x(t-k)$ by integers k. Given this, all subspaces V_j are also shift-invariant. Combining dilation invariance leads to the following conclusion: $x(t) \in V_0 \Rightarrow x(2^j t - k) \in V_j$.

The final requirement is that there exists a function ϕ such that its translates from an orthonormal basis for V_0, i.e., $\{\phi(t-k), k \in \mathbb{Z}\}$ is a basis for V_0. $\phi(2t-k)$ is an orthonormal basis for V_1 by scale invariance, $\{\phi(2t-k), k \in \mathbb{Z}\}$. Similarly, $\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k)$ forms an orthonormal basis for V_j. The function ϕ, which generates the basis functions for all the spaces, $\{V_j\}$, is called the scaling function of MRA.

In summary, a multi-resolution analysis of $\mathbb{L}^2(\mathbb{R})$ is a nested sequence of subspaces $\{V_j\}_{j\in\mathbb{Z}}$ ($\mathbb{Z}$ is the set of integers) such that

- $... \subset V_{-1} \subset V_0 \subset V_1 \subset ... \subset \mathbb{L}^2(\mathbb{R})$

- $\cap_j V_j = \{0\}$, $\overline{\cup_j V_j} = \mathbb{L}^2(\mathbb{R})$

- $x(t) \in V_j \Longleftrightarrow x(2t) \in V_{j+1}$

- $x(t) \in V_0 \Longrightarrow x(t-k) \in V_0$

- $\exists \quad \phi(t)$, such that $\{\phi(t-k)\}$ is an orthonormal basis of V_0.

The literature on MRA is comprehensive. See [48], [25], [43], [1], and [24].

The definition of MRA provides a method of decomposing a function $x(t)$ into a smooth part plus details. At resolution level j, $x(t)$ is approximated by $x_j(t)$, therefore, $x_j(t) \in V_j$. The details $d_j(t)$ are in W_j. At the next level of resolution, $j+1$, the approximation to $x(t)$ is $x_{j+1}(t)$, which includes the details $d_j(t)$ at resolution level j; therefore, $x_{j+1}(t) = x_j(t) + d_j(t)$. The original function $x(t)$ is recovered when the resolution goes to infinity:

$$x(t) = x_j(t) + \sum_{i=j}^{\infty} d_i(t). \tag{22}$$

3.2.3 Wavelet Transform

A wavelet is a waveform bounded in both frequency and time and used in representing data or other functions, the same idea as that used in the FT. However, in wavelet analysis, the fundamental idea is to analyze according to scale, which plays a special role in data analysis. Wavelet analysis processes data at different scales or resolutions. Gross features of data can be obtained through a large "window" and fine features through a small "window." The result of wavelet analysis is to see both the "forest" and the "trees" [41].

The WT solves the dilemma of resolution to a certain extent. Figure 25 is commonly used to explain how time and frequency resolution can be interpreted. Every box in Figure 25 corresponds to a value of the WT in the time-frequency plane. The non-zero area of each box implies that the value of a particular point in the time-frequency plane cannot be known. The areas of all the boxes are the same and determined by Heisenberg's principle, but the widths and lengths can change in WT, representing different proportions of time and frequency. The lower the frequencies, the longer the width of the boxes; and the better the frequency resolution, the poor time resolution, and vice versa. This is how WT deals with the resolution problem. For more information about this topic, see [34] and [1].

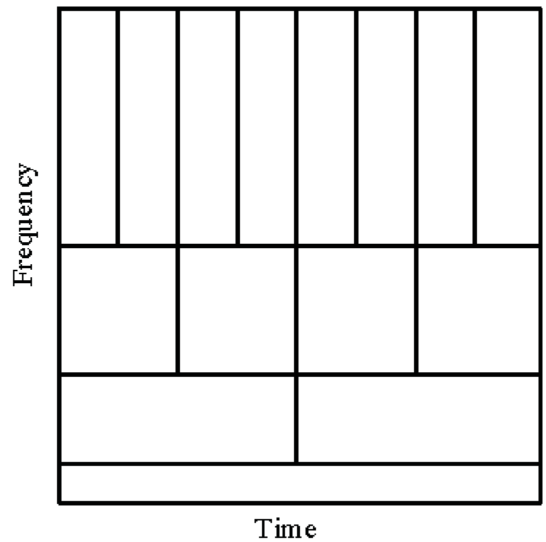


Figure 25: Frequency-Time Domain of Wavelet Transform

The basic difference between WT and FT is in the basis functions used for the transforms. Wavelet functions are localized in space, while the basis functions for FT, sine and cosine, are not. This localization makes wavelets well-suited for approximating data with sharp spikes or discontinuities. WT, unlike FT which utilizes just the sine and cosine as basis functions, does not have a single set of basis functions. WT utilizes an infinite sets of possible basis functions. Therefore, WT can provide information that can be obscured by other time-frequency methods, such as Fourier analysis.

The time-frequency resolution differences between the FT and WT can be illustrated by Figure 26, which shows the basis function coverage of the time-frequency domain for FT and WT, respectively. The left graph shows a Windowed Fourier transform, in which the window is simply a square wave obtained by truncating the sine or cosine function so that it fits a window of a particular width. The resolution of the analysis is the same at all locations in the time-frequency plane because of the use of a single window for all frequencies in the STFT. The right graph shows that WT utilizes various windows for analysis. Short windows are appropriate for isolating discontinuities in the data series, and long basis functions are appropriate for obtaining detailed frequency analysis.

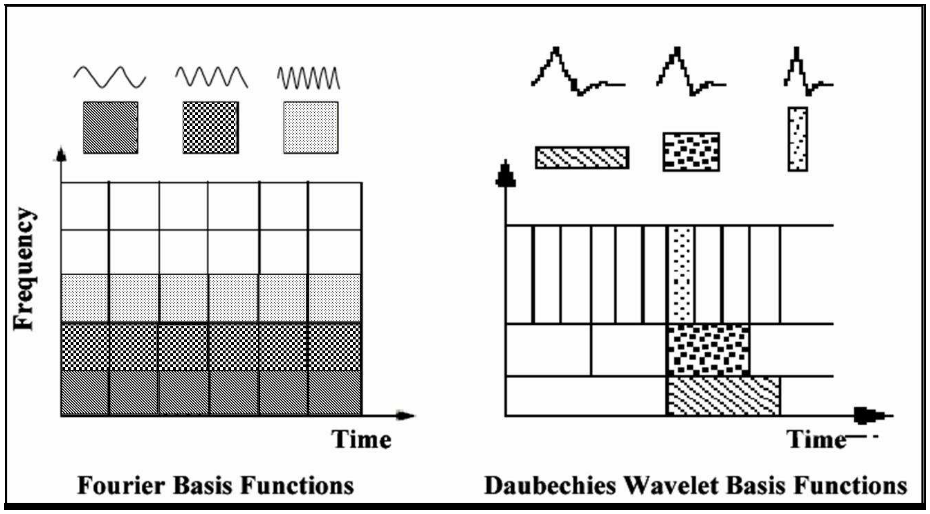


Figure 26: Time-Frequency Tiles and Coverage of the Time-Frequency Plane

To overcome the resolution problem, continuous wavelet transform (CWT) was developed as an alternative approach to the STFT. It is done by multiplying the data by wavelets and computing them separately for different segments of the time-domain data. In this sense, the CWT is similar to STFT with the wavelets replacing the window function. However, they differ in two ways:

1. Data series with sharp discontinuities will be seen when wavelets are used as the window function.

2. The width of the window is changed as the transform is computed for every single spectral component.

3.2.3.1 Continuous Wavelet Transform

Let $f(t)$ and $g(t)$ be two functions in $\mathbb{L}^2[a,b]$. The inner product of the two functions is defined by Equation (23)

$$\langle f(t), g(t)\rangle = \int_a^b f(t)g^*(t)dt. \tag{23}$$

Based on the concept of the inner production of functions, the CWT is defined as the inner product of the signal $x(t)$ with the basis function $\psi_{\tau,s}(t)$:

$$CWT_x^{\psi}(\tau, s) = \langle x(t), \psi_{\tau,s}(t)\rangle = \int x(t)\psi^*_{\tau,s}(t) = \frac{1}{\sqrt{|\,s\,|}}\int x(t)\psi^*(\frac{t-\tau}{s})dt, \tag{24}$$

where

$$\psi_{\tau,s} = \frac{1}{\sqrt{s}}\psi(\frac{t-\tau}{s}). \tag{25}$$

The transformed data series is a function of two variables, the translation parameter τ and the scale parameter s. Translation is related to the location of the window, as the window shifts throughout the data. It corresponds to the time information in the transform domain. Scale is defined as the inverse of frequency. Large scales correspond to the global feature of a data series that usually spans the entire data, whereas small scales correspond to detailed information of a hidden pattern in a data series that usually lasts a relatively short time. Scaling works as a mathematical operation that either dilates (large scales) or compresses (small scales) a data series.

$\psi(t)$ is the transforming function, called the mother wavelet. The term "mother wavelet" derives from some important properties of the wavelet analysis:

- "Wavelet" means a small wave function.
  - "Small" refers to that the support of the function is short and small.
  - "Wave" refers to the condition that this function is oscillatory.
- "Mother" implies that the functions with different regions of support used in the transform are derived from one main function.

CWT is reversible, provided that the admissibility condition in Equation (26) is satisfied,

$$c_\psi = \{2\pi \int_{-\infty}^{\infty} \frac{|\hat{\psi}(\xi)|^2}{|\xi|}\}^{1/2} < \infty, \tag{26}$$

where c_ψ is the admissibility constant, which depends on the wavelet used. $\hat{\psi}(\xi)$ is the FT of $\psi(t)$. The admissibility condition implies that $\hat{\psi}(0) = 0$, which is

$$\int \psi(t)dt = 0. \tag{27}$$

Equation (27) is not a very restrictive requirement since many wavelet functions whose integral is zero can be found. Then the reconstruction is realized through

$$x(t) = \frac{1}{c_\psi^2}\int_s \int_\tau CWT_x^\psi(\tau,s)\frac{1}{s^2}\psi(\frac{t-\tau}{s}). \tag{28}$$

Literature on CWT is comprehensive. See [48], [59], and [1].

3.2.3.2 Discretized Continuous Wavelet Transform

CWT cannot be easily be computed with analytical equations. So the transforms must be discretized. The most intuitive way to do this is to sample the time-scale plane with a sampling rate, which depends on the scale. According to Nyquist's rule, if the time-scale plane is sampled with a sampling rate of r_1 at scale s_1, the sampling rate at scale s_2 is proportional to the ratio of scales $r_2 = \frac{s_1}{s_2} r_1$. At higher scales, the sampling rate can be decreased and will thus save a considerable amount of computation time. Nyquist's sampling rate is the minimum sampling rate that allows the original continuous data to be reconstructed from its discrete samples.

The discretization procedure can be expressed as follows: Let scale discretization be $s = s_0^j$, and translation discretization be $\tau = k \cdot s_0^j \cdot \tau_0$ with $s_0 > 1$ and $\tau_0 > 0$. Inserting these two terms into the CWT Equation (25) renders

$$\psi_{j,k}(t) = s_0^{-j/2} \psi(s_0^{-j} t - k\tau_0). \tag{29}$$

With $\psi_{j,k}$ constituting an orthonormal basis, the wavelet series transform becomes

$$DCWT_x^{\psi}(\tau, s) = \int x(t) \psi_{j,k}^*(t) dt, \tag{30}$$

or

$$x(t) = c_\psi \sum_j \sum_k DCWT_x^{\psi}(\tau, s) \psi_{j,k}(t). \tag{31}$$

3.2.3.3 Discrete Wavelet Transform

The discretized continuous wavelet transform (DCWT) enables the computation of the CWT through sampling. The information provided by the DCWT is highly redundant for the purpose of reconstructing the data series. The discrete wavelet transform (DWT) provides sufficient information both for the analysis and the synthesis of the data series, but with a significant reduction in computation time. Therefore, the DWT is considerably easier to implement than the DCWT.

The DWT utilizes filters with different cutoff frequencies to analyze the data at different scales. Filtering is an operation that maps from $\mathbb{L}^2(\mathbb{Z})$ to $\mathbb{L}^2(\mathbb{Z})$. With $\mathbf{H}$ denoting the filter, for $x \in \mathbb{L}^2$, $y = \mathbf{H}x$ has a component-wise representation

$$y(n) = (h * x)(n) = \sum_k h(k)x(n-k),$$

where $h(k) = h_k, \quad k \in \mathbb{Z}$ are filter coefficients that are obtained when filter $\mathbf{H}$ is applied to the unit impulse function at zero $u = \{\dots, 0, 0, 1, 0, 0, \dots\}$, so that

$$h = \mathbf{H}u = \{\dots, h_0, h_1, \dots\}.$$

Low pass filter $\mathbf{H}$ is used to average or smooth the data series, with low frequencies preserved. High pass filter $\mathbf{G}$ is to difference the data series, with high frequencies preserved. Therefore, the analysis of high frequencies is enabled by passing the data series through a series of high-pass filters, and the analysis of low frequencies by passing the data series through a series of low-pass filters. Filtering a data series is also expressed as the mathematical operation of the convolution of the data with the impulse response of the filter:

$$\mathbf{H} : \mathbb{L}^2(\mathbb{Z}) \mapsto \mathbb{L}^2(\mathbb{Z}) \quad (\mathbf{H}a)_k = \sum_n h_{n-k}a_n,$$

$$\mathbf{G} : \mathbb{L}^2(\mathbb{Z}) \mapsto \mathbb{L}^2(\mathbb{Z}) \quad (\mathbf{G}a)_k = \sum_n g_{n-k}a_n.$$

Filtering operations change the resolution of the data series, which is related to the amount of detail information in the data. A half-band low-pass filter removes all the frequencies that are above half of the highest frequency in the data and halves the resolution, which can be interpreted as a loss of half of the information.

The scale is changed by decimation $[\downarrow 2]$ and dilation $[\uparrow 2]$ operations, which leave the resolution unchanged. Decimation maps from $\mathbb{L}^2(\mathbb{Z})$ to $\mathbb{L}^2(2\mathbb{Z})$ defined as

$$([\downarrow 2]x)_k = \sum_n x_n \delta_{n-2k} = x_{2k},$$

where δ is the Dirac function. When decimation is applied to a data series, only the values on positions with even indices are retained. This process halves the number of points and doubles the scale of the data series. Decimation (or dilation) after low-pass filtering (or high-pass filtering) will not change the resolution. If half of the spectral components are removed by low-pass filtering, half of the number of samples are redundant. Therefore, halving the samples by decimation does not lose any information. Dilation maps from $\mathbb{L}^2(2\mathbb{Z})$ to $\mathbb{L}^2(\mathbb{Z})$ defined as

$$([\uparrow 2]x)_k = \sum_n x_n \delta_{k-2n}.$$

When dilation is applied to a data series, zeroes are inserted between the original values to expand it. This process increases the sampling rate by adding new samples to the data series.

Let $\mathcal{H} \equiv [\downarrow 2]\mathbf{H}$ and $\mathcal{G} \equiv [\downarrow 2]\mathbf{G}$, and the data series $x = \{x_n\}$:

$$(\mathcal{H}x)_k = \sum_n h_{n-2k} x_n, \tag{32}$$

$$(\mathcal{G}x)_k = \sum_n g_{n-2k} x_n. \tag{33}$$

An application of operators $\mathcal{H}$ and $\mathcal{G}$ corresponds to one step in the DWT. Denote the original data by $X = x^{(J)} = \{x_k^{(J)}\}$, which has a length 2^J. This process,

$$X \mapsto (\mathcal{H}^k X, \mathcal{G}\mathcal{H}^{k-1} X, \ldots, \mathcal{G}\mathcal{H}^2 X, \mathcal{G}\mathcal{H}X, \mathcal{G}X),$$

is also illustrated in Figure 27.

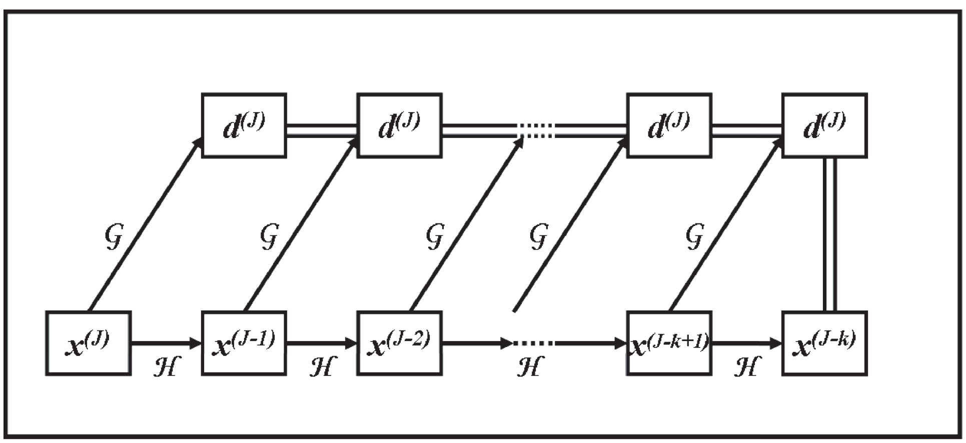


Figure 27: Decomposition Algorithm

Each step of the DWT moves the data series to the next coarser approximation (level) $x^{(j-1)}$ by applying $\mathcal{H}$, $x^{(j-1)} = \mathcal{H}x^{(j)}$. The detail information, lost by approximating $x^{(j)}$ by the "averaged" $x^{(j-1)}$, is given by $d^{(j-1)} = \mathcal{G}x^{(j)}$. Hence, the DWT of the data series X of length 2^J can be be represented as

$$(x^{(J-k)}, d^{(J-k)}, d^{(J-k+1)}, d^{(J-k+2)}, \ldots, d^{(J-2)}, d^{(J-1)}). \tag{34}$$

The decomposition process has several important characteristics:

- The time resolution is halved since the data series is only characterized by half the number of samples. The frequency resolution is doubled since the frequency band spans only half the previous frequency band, thus reducing the uncertainty in the frequency by half.
- The time localization of frequencies will not be lost. The most prominent frequencies in the original signal have high amplitude in the region of the DWT data series in which they are included.
- The time localization has a resolution that depends on the level on which it appears. The time localization of the information contained in high frequencies is more precise because it is characterized by more number of samples. Thus, the process offers a good time resolution at high frequencies and good frequency resolution at low frequencies.

DWT is a reversible process. By defining operators $\mathcal{H}^*$ and $\mathcal{G}^*$ as follows:

$$(\mathcal{H}^* x)_n = \sum_k h_{n-2k} x_k, \tag{35}$$

$$(\mathcal{G}^* x)_n = \sum_k g_{n-2k} x_k. \tag{36}$$

$x^{(j)}$ can be reconstructed as

$$x^{(j)} = \mathcal{H}^* x^{(j-1)} + \mathcal{G}^* d^{(j-1)} = \mathcal{R}(x^{(j-1)}, d^{(j-1)}). \tag{37}$$

Recursive application of Equation (37) leads to

$$\sum_{i=1}^{k-1} (\mathcal{H}^*)^{k-1-i} \mathcal{G}^* d^{(J-k+i)} + (\mathcal{H}^*)^k x^{(J-k)} \mapsto X. \tag{38}$$

This process is illustrated in Figure 28.

See [48], [59], and [1] for more information on the DWT.

3.2.3.4 Non-Decimated Wavelet Transform

The two main types of wavelet transforms are continuous and discrete wavelet transforms. DWT is very efficient from the computational point of view. One intrinsic property of the

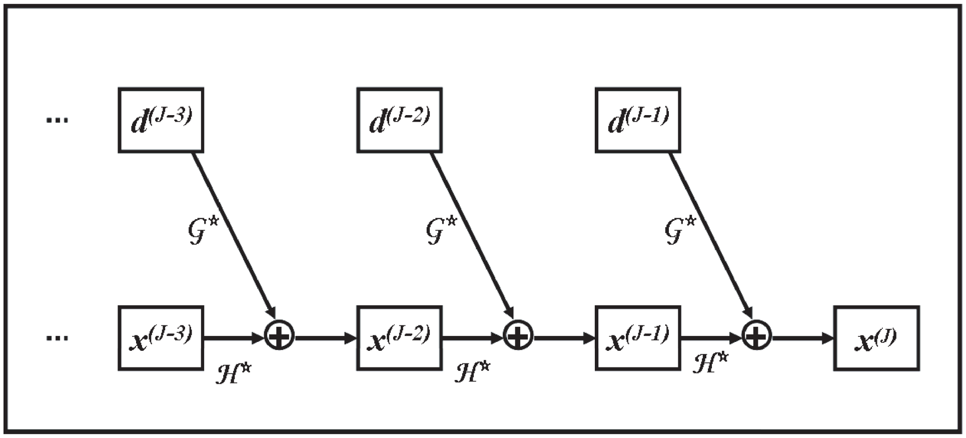


Figure 28: Reconstruction Algorithm

DWT is the decimation of the wavelet coefficients, which removes every other coefficients of the current level. Thus, the WT can be done in a fast and compact manner, and the inverse transform can be perfectly done from the remaining coefficients. Unfortunately, the decimation results in a translation variance of the WT. The translation variance means that the DWTs of a data series and its translations are not the same (see Figure 29). A data series is shown in the upper left graph, and its right translation is shown in the upper right graph. The graphs on the bottom show the DWTs of the data series in the graphs above them. It is obvious the DWT coefficients are too different to be obtained by shifting the other one.

Non-decimated wavelet transform (NDWT), however, does not decimate the data series. It gives an increasing amount of information that can be used to obtain more accurate and comprehensive understanding of data series properties. The number of wavelet coefficients does not shrink between the transformed levels. Due to the redundance in the coefficients, NDWT has larger storage requirements and involves more computations.

Let S^k: $\mathbb{L}^2(\mathbb{Z}) \mapsto \mathbb{L}^2(\mathbb{Z})$ be the shift operator defined coordinate-wise as

$$(S^k x)_n = x_{n+k},$$

and let

$$\mathcal{D}_0 = [\downarrow 2] \text{ and } \mathcal{D}_1 = [\downarrow 2]S$$

be a pair of decimation operators that decimate by retaining values at even and odd indices.

The operator $\mathcal{D}_0$ was used earlier in the DWT. A single step in the DWT was defined as an action of filters $\mathbf{H}$ and $\mathbf{G}$ followed by decimation

$$x^{(j-1)} = \mathcal{D}_0 \mathbf{H} x^{(j)} \text{ and } d^{(j-1)} = \mathcal{D}_0 \mathbf{G} x^{(j)}.$$

The reconstruction step was

$$x^{(j)} = \mathcal{R}(x^{(j-1)}, d^{(j-1)}) = \mathcal{R}_0(x^{(j-1)}, d^{(j-1)}).$$

An orthogonal decomposition can be obtained by applying $\mathcal{D}_1$

$$x_1^{(j-1)} = \mathcal{D}_1 \mathbf{H} x^{(j)} \text{ and } d_1^{(j-1)} = \mathcal{D}_1 \mathbf{G} x^{(j)},$$

with the reconstruction step

$$x^{(j)} = \mathcal{R}_1(x_1^{(j-1)}, d_1^{(j-1)}).$$

Vectors $x_1^{(j-1)}$ and $d_1^{(j-1)}$ are different from $x^{(j-1)}$ and $d^{(j-1)}$, but the underlying transform is still orthogonal.

For quadrature mirror filters h and g, define dilation filters $h^{[r]}$ and $g^{[r]}$ in a recursive way:

$$h^{[0]} = h, \quad g^{[0]} = g,$$

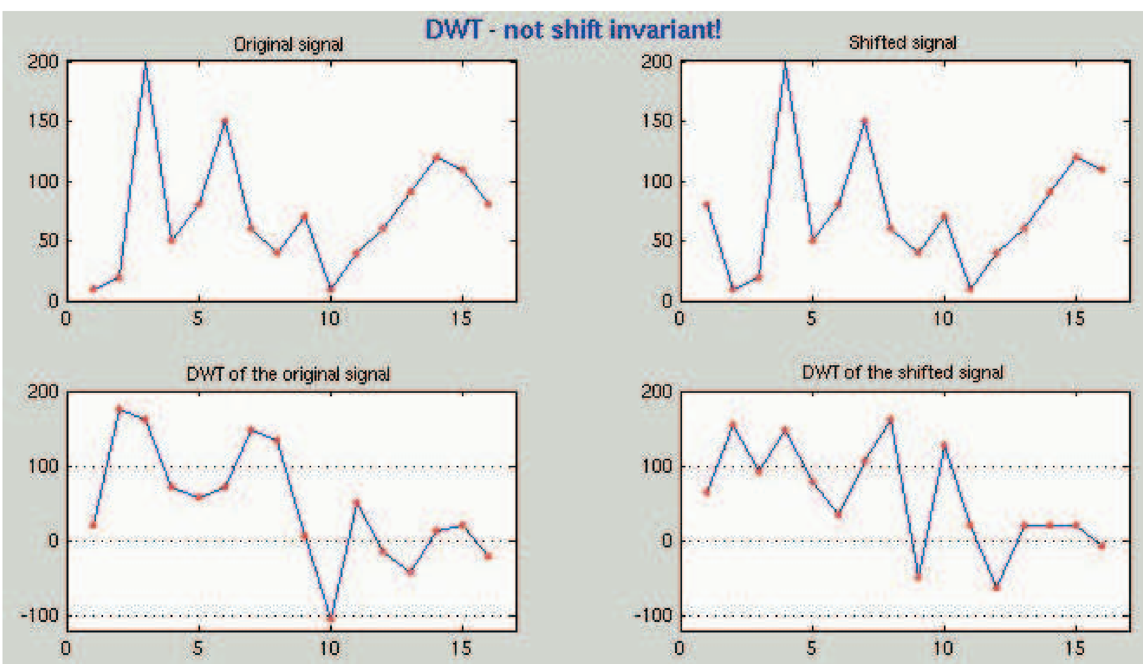


Figure 29: The Discrete Wavelet Transform Lacks Translation-Invariance

$$h^{[r]} = [\uparrow 2]h^{[r-1]}, \quad g^{[r]} = [\uparrow 2]g^{[r-1]}.$$

Let $H^{[r]}$ and $G^{[r]}$ be convolution operators with filters $h^{[r]}$ and $g^{[r]}$, respectively. A non-decimated WT is defined as a sequential application of operators (convolutions) $H^{[j]}$ and $G^{[j]}$ to a given data series. The process can be expressed as

$$x^{(j-1)} = \mathbf{H}^{[J-j]}x^{(j)}$$

$$d^{(j-1)} = \mathbf{G}^{[J-j]}x^{(j)}.$$

The non-decimated WT of $x^{(J)}$ is a vector

$$(d^{(J-1)}, \quad d^{(J-2)}, \quad , \ldots, d^{(J-j)}, x^{(J-j)}),$$

for some $j \in \{1, 2, \ldots, J\}$, representing the depth of the transform. Subvectors $d^{(J-1)}$, $d^{(J-2)}$, $\ldots$, $d^{(J-j)}$ are levels of detail while the subvector $x^{(J-j)}$ is the "smooth."

The operator $(\mathbf{H}^{[j]}, \mathbf{G}^{[j]})$ is not orthogonal but $(\mathcal{D}_0\mathbf{H}^{[j]}, \mathcal{D}_0\mathbf{G}^{[j]})$ and $(\mathcal{D}_1\mathbf{H}^{[j]}, \mathcal{D}_1\mathbf{G}^{[j]})$ are each orthogonal. The first pair of transforms produces values at even indices in $x^{(J-j-1)}$ and $d^{(J-j-1)}$, and the second produces the values at odd indices. Let $\mathcal{R}_0^{[j]}$ and $\mathcal{R}_1^{[j]}$ be their inverse transforms. Then,

$$x^{(j)} = \overline{\mathcal{R}}^{[J-j]}(x^{(j-1)}, d^{(j-1)}),$$

for $\overline{\mathcal{R}}^{[j]} = (\mathcal{R}_0^{[j]} + \mathcal{R}_1^{[j]})/2$. For more information on non-decimated wavelet transform, see [104] and [72].

The following example shows the differences between DWT and NDWT. Figure 30 presents the data series: Doppler. Figure 31 shows the coefficient distribution of the DWT, and Figure 32 shows the differences between DWT and NDWT.

3.2.4 Wavelet Families

A number of basis functions can be used as the mother wavelet for the WT. The characteristics of the resulting WT are determined by the mother wavelet because it produces all the wavelets used in the transform through translation and scaling. Therefore, the details of the particular application should be taken into account, and the appropriate mother wavelet should be chosen in order to use the WT effectively.

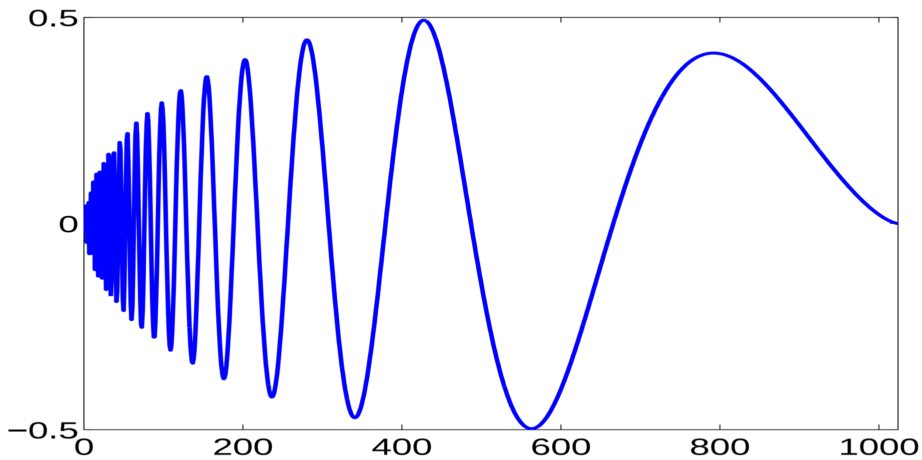


Figure 30: Data Series: Doppler

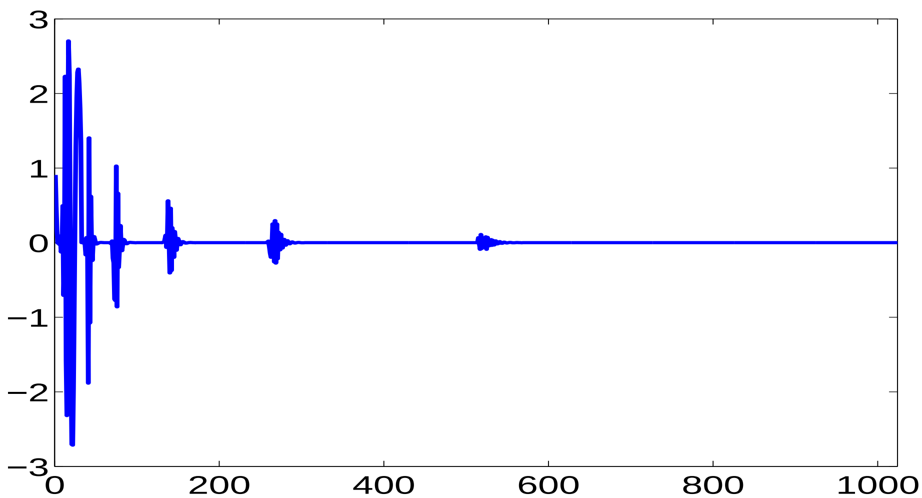


Figure 31: Doppler in the Wavelet Domain through the DWT

Most wavelets do not have explicit formulas, but some wavelets such as the Haar wavelet do. See Figure 33(a)

$$\omega(t) = \begin{cases} 1 & \text{if } 0 \leq t \leq 1/2 \\ -1 & \text{if } 1/2 \leq t \leq 1 \\ 0 & \text{otherwise} \end{cases} \tag{39}$$

Orthogonal and compactly supported wavelets include Daubechies, Symmlets, Coiflets, and so forth, but with implicit formulas for ϕ and ψ.

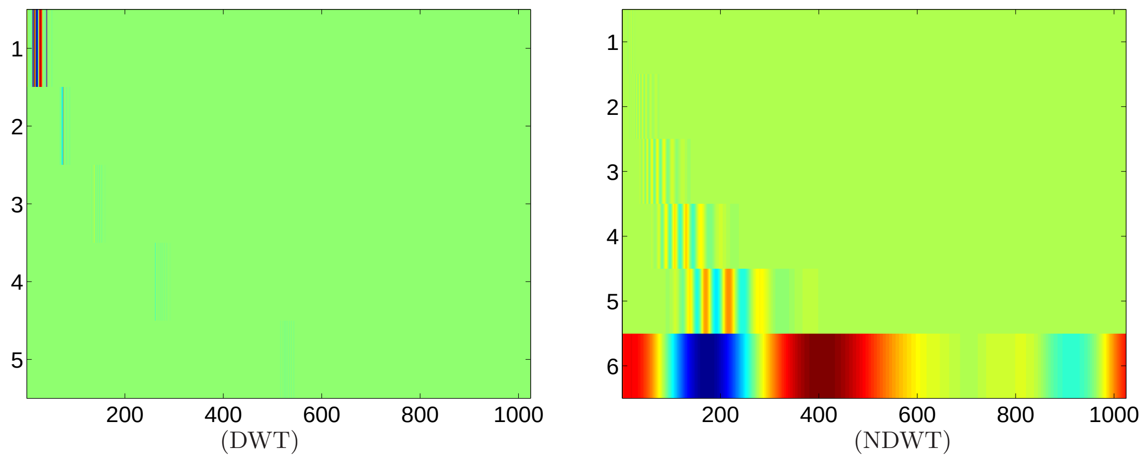


Figure 32: Wavelet Transform by the DWT and the NDWT

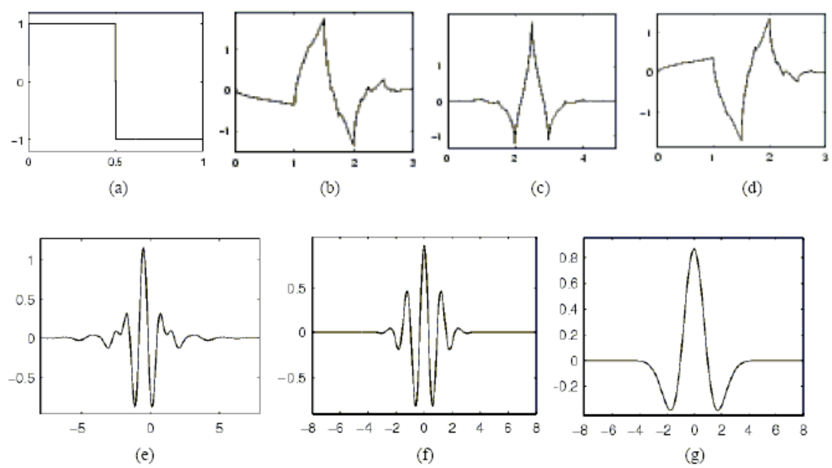


Figure 33: Wavelet Families (a) Haar (b) Daubechies4 (c) Coiflet1 (d) Symmlet2 (e) Meyer (f) Morlet (g) Mexican Hat

3.3 Forecasting Method - WAW

3.3.1 Forecasting Methodology

Wavelet decomposition is a relatively novel methodology developed in the last two decades. The wavelet domain and more generally multi-scale domains, are especially suitable for modeling time series. Wavelet-based representations of time series describe how time series evolve over time at a given scale that is either an interval (span of time) or a spatial area. Wavelets are atomic functions that are compactly supported and integrated to zero, and waving above and below the x-axis. As such, wavelets are building blocks that are suitable for the localized phenomena of varying frequencies.

Among the host of various WTs, the non-decimated (stationary, translation invariant, maximum overlap) transform is most suitable for tasks of forecasting. The standard orthogonal WT are most parsimonious but lack the shift invariance and are calculationally unsuitable for time series forecasting, see Figure 34. If the observations in the time series are equispaced, WTs are extremely fast (faster than FFT) and computationally amount to a filtering problem. The implementational difference between standard orthogonal WTs and NDWTs is the way in which filtering is applied. NDWTs use filtering without down sampling, producing redundant but shift-invariant decompositions.

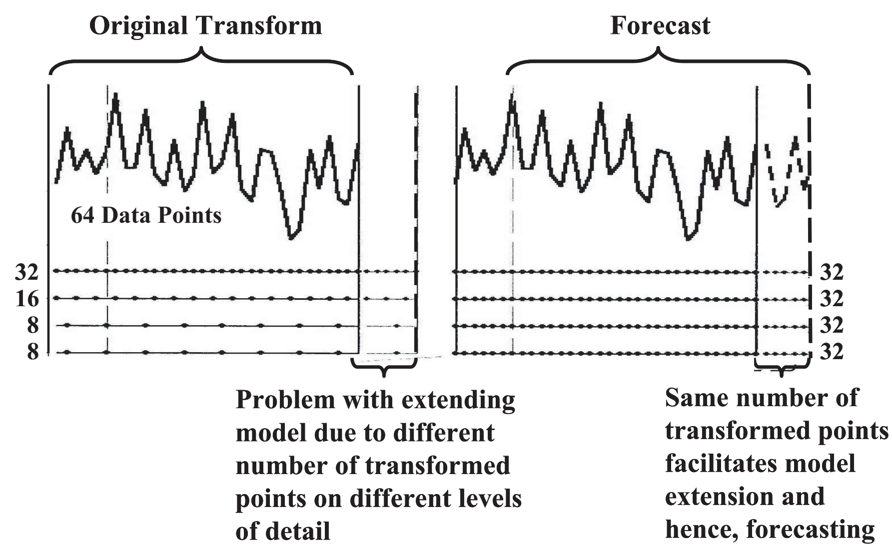


Figure 34: Decimated and Non-decimated Wavelet Transforms

One characteristic of wavelets is their multi-scale filtering, which facilitates the separating of a data series into various levels of scales that describe the details in various resolutions. This ability is utilized to "zoom in" at particular time scales to de-trend and de-seasonalize a time series. The trend component is "located" in scaling coefficients and on coarse levels of detail (lower frequencies) as opposed to the high-frequency component, which requires fine-grained detail space for its description. The signature of the seasonal component is located at the intermediate levels. In this manner, by separating coarse, intermediate, and fine levels of detail, the time series may be de-trended, de-seasonalized, and de-noised in a mathematically logical way.

For each level, a suitable technique for analyzing the data and making predictions is found. The main processes of modeling techniques, Auto-Regressive Moving Average with external input (ARMAX) model, harmonic regression, and Holt-Winters' method are addressed first.

The Auto-Regressive Moving Average (ARMA) model is a static time series model applicable to a time series with neither trends nor seasonality that exhibit time homogeneity. The ARMAX model is a generalization of the ARMA model, which is capable of incorporating an external, (X), input variable. The form of the ARMAX model is

$$\Phi(B)y_t = \Xi(B)x_{t-\alpha} + \Theta(B)\varepsilon_t,$$

where $x_{t-\alpha}$ is an external input variable, y_t is the response (output variable), ε_t is the white noise, α is the lag delay between the input and the output, and B is the backshift operator. The polynomials in the backshift operator Φ, Ξ, and Θ are given by

$$\begin{aligned} \Phi(B) &= 1 + \phi_1 B + \phi_2 B^2 + ... + \phi_{n\phi} B^{m_\phi}, \\ \Xi(B) &= 1 + \xi_1 B + \xi_2 B^2 + ... + \xi_{n\xi} B^{m_\xi}, \quad \text{and} \\ \Theta(B) &= 1 + \theta_1 B + \theta_2 B^2 + ... + \theta_{n\theta} B^{m_\theta}. \end{aligned}$$

Literature on the ARMAX model and its generalization is comprehensive [70]. In the proposed methodology, the ARMAX model will be utilized to account for the external business environment, so the forecasting method does not solely depend on historical data.

Harmonic regression (trigonometric regression, cosinor regression) is a linear regression model in which the predictor variables are trigonometric functions of a single variable, usually a time-related variable. Harmonic regression is utilized in phenomena that tend to exhibit periodic behavior.

A simple harmonic regression model is

$$Y = \alpha_0 + \sum_{n=1}^{N}(\beta_n \cos(n\omega x) + \gamma_n \sin(n\omega x)),$$

where ω is the frequency.

More general models are

$$Y = \alpha_0(x) + \sum_{n=1}^{N}(\beta_n(x) \cos(n\omega x) + \gamma_n(x) \sin(n\omega x)),$$

where the values of α, β, and γ depend on x. See [45] for more utilization of harmonic regression.

Holt-Winters' seasonal method [16] is an approach that applies to time series containing both trend and seasonal variations. Holt-Winters' method, which does not assume any stochastic structure of a time series, is based on three smoothing equations. The method is as follows:

If the observed time series $Y_1, Y_2, ...Y_n$ contains not only the trend, but seasonality with period d as well, then the forecasting function that takes them into account is

$$P_n Y_{n+h} = \hat{a}_n + \hat{b}_n h + \hat{c}_{n+h}, \quad h = 1, 2, \ldots,$$

where $\hat{a}_n, \hat{b}_n,$ and $\hat{c}_n$ are the estimates of the trend level, trend slope, and seasonal component at time n, respectively:

$$\begin{aligned}
\hat{a}_{n+1} &= \alpha(Y_{n+1} - \hat{c}_{n+1-d} + (1-\alpha)(\hat{a}_n + \hat{b}_n), \\
\hat{b}_{n+1} &= \beta(\hat{a}_{n+1} - \hat{a}_n) + (1-\beta)\hat{b}_n, \\
\hat{c}_{n+1} &= \gamma(Y_{n+1} - \hat{a}_{n+1}) + (1-\gamma)\hat{c}_{n+1-d}, \quad \text{and} \\
\hat{c}_{n+h} &= \hat{c}_{n+h-kd}, \ h = 1, 2, \ldots, \quad \text{with} \quad n + h - kd \leq n.
\end{aligned}$$

The initial conditions are

$$\begin{aligned}
\hat{a}_{d+1} &= Y_{d+1}, \\
\hat{b}_{d+1} &= (Y_{d+1} - Y_1)/d, \\
\hat{c}_i &= Y_i - (Y_1 + \hat{b}_{d+1}(i-1)), \ i = 1, \ldots, d+1,
\end{aligned}$$

and α, β, and γ are preset parameters. More on Holt-Winters' method is available in [16], [52] and [79].

The de-trended and de-seasonalized time series by NDWT should have a stationary signature. Hence, the ARMA part of an ARMAX model should be able to describe this stationary high-frequency component, and at the same time, the input of the ARMAX model will enable the model to take into account external inputs. Thus, the high-frequency component filtered out by the wavelet technique can be fitted by an ARMAX model, which will be used to make forecasts for the high frequency component in the sequel. For the trend and seasonal components represented by wavelet coefficients at various levels of detail (at various frequencies), predictions will be made for future observations. All predictions are done in the wavelet domain. Subsequently, the predicted values for the trend, seasonality, and high-frequency components will be combined via the inverse wavelet transform to obtain the final forecast.

The forecasting process is summarized as follows:

1. Apply the NDWT to the historical time series to separate the trend and seasonal components from the high-frequency component. This separation is done in the wavelet domain.

2. Predict the future value of the trend using Holt-Winters' method. This prediction is done on the "smooth" part of wavelet decomposition, i.e., on the scaling coefficients.

3. Predict the seasonality component using harmonic regression with estimated seasonal periods.

4. Apply the ARMAX model to predict the high-frequency component.

5. Combine the predicted trend, seasonality and high-frequency component to obtain the required forecast. This step involves the inverse wavelet transform of the predicted values at different scales.

The wavelet-ARMAX-HoltWinter model can be applied to three sets of data: historical data for the natural gas electricity purchase prices (cnt/mcf), the residential and commercial consumption of electricity (Tbtu), and the electricity industrial sector prices (hcnt/kwh).

3.3.2 Forecasting Error Analysis

For any forecasting method, modeling errors are unavoidable. The behavior of the modeling errors during the WT and the inverse WT might have a significant impact on the accuracy of the forecasting. The WAW method utilizes the NDWT to separate historical data into various levels of scale, and then analyzes each level at a resolution matched to its scale.

Prediction is done in the wavelet domain. By inverse WT, a prediction in the time domain is obtained. The exact statistical analysis is possible in principle since the transform and the models used at various scales are known, but their simultaneous treatment is overly complex. This necessitates the understanding of the behavior of the modeling errors when the WT and inverse WT are done to a time series [19].

The behavior of these errors are investigated using simulation techniques. Several scenarios have been chosen to explore the behavior of the WAW procedure. Such simulations result in useful and informational empirical analyses since the inputs are controlled.

- **Experiment 1**

  In the time domain, white standard normal noise is transformed to the wavelet domain by NDWT using the Symmlet 8 filter. The Gaussianity is tested at each level of scales in the wavelet domain. Figure 35 shows the $Q-Q$ plot at each level. Figure 35 clearly illustrates that white standard normal noise on the input was transformed to the levels that looked marginally normal.

  The energies for the time series in the time domain and at each level in the wavelet domain are shown in Table 2. The goal of this analysis was to investigate the propagation of the error energies in the wavelet domain. It is concluded that the errors at each level in the wavelet domain have a magnitude comparable to those of the input data set in the time domain.

Table 2: Energy at Each Level and the Recovered data

L	TS	1st L	2nd L	3rd L	4th L	5th L	6th L	7th L	8th L
E	1012.2	1050.6	938.8	1154.6	961.9	1106.2	1091.6	1017.8	853.8

- **Experiment 2**

  In the wavelet domain, assign each level white standard normal noise. When the errors at each level are combined by inverse transform to the time domain using the Symmlet 8 wavelet filter, the recovered data in the time domain preserves normality (see Figure 36).

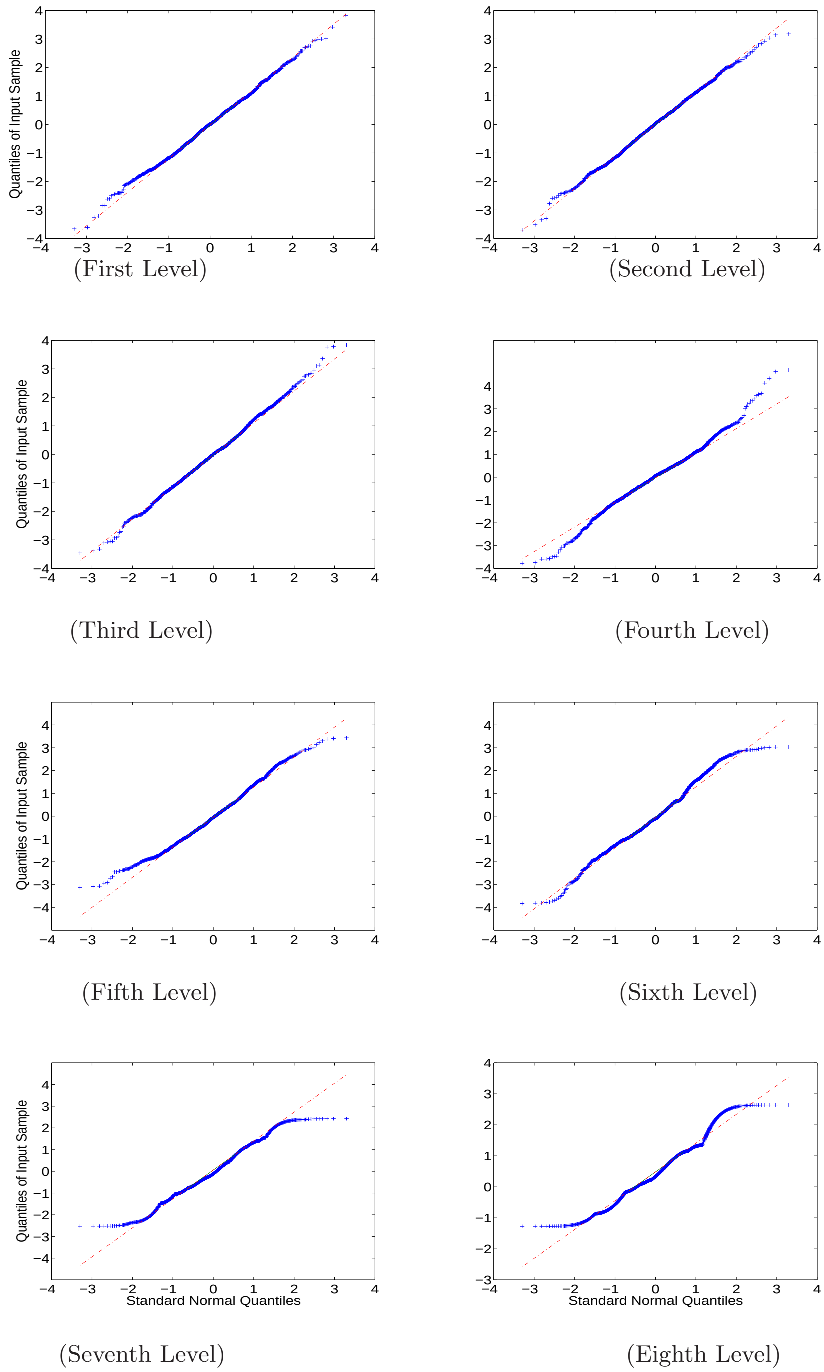


Figure 35: QQ Plot of Sample Data versus Standard Normal

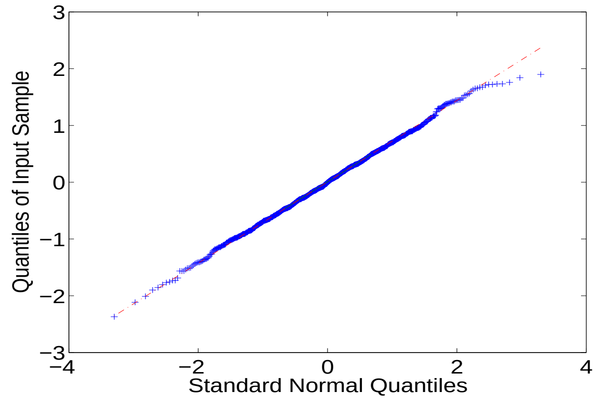


Figure 36: QQ Plot of Sample Data versus Standard Normal

The three tests for the whiteness, Portmanteau, Ljung-Box, and McLeod-Li tests, are performed on the error in the time domain. The resulting p-values suggest that the error is not white noise any more (Table 3). Thus, the inverse NDWT introduces color into the noise.

Table 3: Tests for White Noise

Tests	p-Value
Portmanteau	0.00031799
Ljung-Bbox	0
McLeod-Li	0

The energies for the errors at each level in the wavelet domain and for those of the recovered data in the time domain are shown in Table 4. This experiment shows that the errors of the recovered data in the time domain have a magnitude comparable to those at each level in the wavelet domain; that is, the errors in the wavelet domain are not magnified when transformed back to the time domain.

Table 4: Energy at Each Level and the Recovered Data

L	1st L	2nd L	3rd L	4th L	5th L	6th L	7th L	8th L	Recd TS
E	1065.6	995.1	1016.9	1003.3	1048.7	1060.2	968.0	0	317.3365

The auto-regressive (AR) process can be used to model the errors and ultimately to derive an additional systematic component from them. The part of the errors represented by the AR process can be fed back to the forecasting model. An AR process $\{X_t\}$ of order p is defined by

$$\Phi(B)X_t = Z_t, \tag{40}$$

where $Z_t \sim WN(0, \sigma^2)$ and B is the backshift operator. The polynomial in the backshift operator Φ is given by

$$\Phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p. \tag{41}$$

The partial autocorrelation function (PACF) of an AR process is the function $\alpha(\cdot)$, defined by the equations

$$\alpha(0) = 1, \text{ and } \alpha(h) = \phi_{hh}, \quad h \geq 1,$$

where ϕ_{hh} is the last component of

$$\phi_h = \Gamma_h^{-1} \gamma_h,$$

where $\Gamma_h = [\gamma(i-j)]_{i,j=1}^{h}$, and $\gamma_h = [\gamma(1), \gamma(2), \dots, \gamma(h)]'$ with $\gamma(h), \quad h = 0, \pm 1, \pm 2, \dots$ the autocovariance function at lag h. The PACF of an AR(p) process is zero for lags greater than p. Therefore, if $\{X_t\}$ is an AR(p) process, then the sample PACF, based on observations, should reflect the properties of the PACF itself. In particular, if the sample PACF is significantly different from zero for $0 \leq h \leq p$ and negligible for $h > p$, it suggests that an AR(p) model might represent the data well. To decide what is meant by "negligible", the knowledge can be used, that for an AR(p) process, the sample PACF values at lags greater than p are approximately independent $N(1, 1/n)$, where n is the number of observations of the random variable. This means that roughly 95% of the sample PACF values beyond lag p should fall within the bounds $\pm 1.96/\sqrt{n}$. A sample PACF satisfying $|\hat{\alpha}(h)| > 1.96/\sqrt{n}$ for $0 \leq h \leq p$ and $|\hat{\alpha}(h)| < 1.96/\sqrt{n}$ for $h \geq p$ can be estimated well by an AR(p) process.

The left figure in Figure 37 shows the sample partial autocorrelation function (PACF) of the AR process together with the bounds $\pm 1.96/\sqrt{n}$. It is easy to read off the order of the AR process $p = 6$.

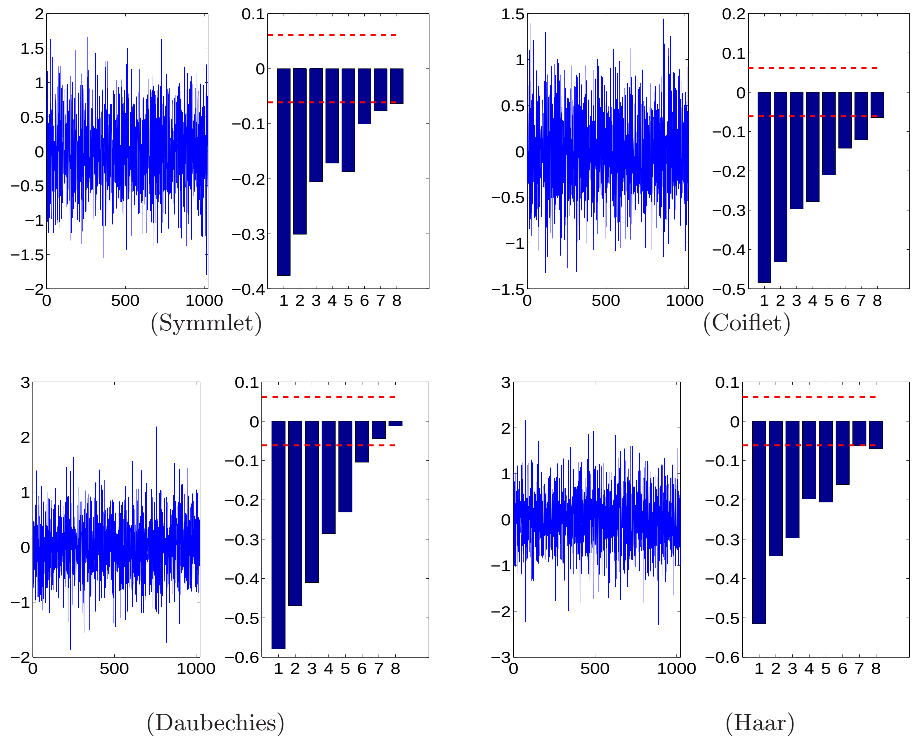


Figure 37: AR Model using Different Wavelet Filters

- **Experiment 3**

  Next, we investigate the robustness of the AR model with respect to the type of wavelet filters. Different wavelet filters are used to assess the impact on the AR process model. Figure 37 shows the resulting AR process for the input data set with length 2^{10} using Symmlet, Coiflet, Daubechies, and Haar wavelet filters. It can be read from the figure that $p = 6$ $p = 7$, $p = 6$, and $p = 6$, respectively. An AR process of order 6 can be used to model the forecasting error. Therefore, the AR model is highly robust regardless of what wavelet filter is used.

- **Experiment 4**

  It is found that the order of the AR process depends on the length of the input data set. If the length of the data set is kept fixed, then the order of the AR process can be determined. However, if the length of the data set changes, then the order of the AR process also requires adjusting. A log-linear relationship is found for the length of data set from 2^5 to 2^{11}: the order of the corresponding AR process is from 1 to 7. The coefficients of these AR processes also exhibit consistency. Figure 38 shows the impact of length on the order of the AR process.

- **Experiment 5**

In the next two experiments, we investigate the impact of the variance of the white noise on the conclusions obtained from the last four experiments.

When the white noise in the time domain is normal but with randomly generated variance, the energies at each level after NDWT are shown in Table 5. Thus, the conclusion from experiment 1, that errors at each level in the wavelet domain have a magnitude comparable to those of the input data set in the time domain, is still valid for white noise with random variance.

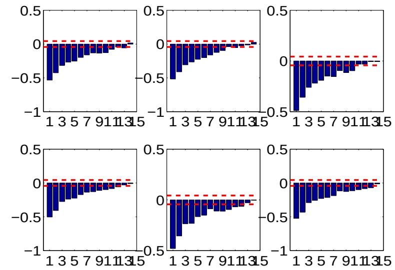
Table 5: Energy at Each Level and the Recovered data

Location	E($\sigma^2 = 1.2621$)	E($\sigma^2 = 0.9656$)	E($\sigma^2 = 0.5185$)
TS	1587.5	951.5	282.23
1st L	2133.0	557.1	332.30
2nd L	1662.6	982.9	286.30
3rd L	1520.2	990.6	276.85
4th L	1571.4	1002.2	256.44
5th L	1123.3	704.4	323.94
6th L	1671.7	553.9	276.98
7th L	1857.4	900.7	241.88
8th L	1485.1	551.6	324.2

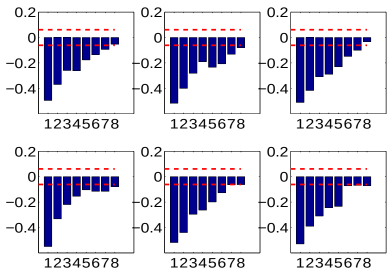
- **Experiment 6**

When the variance of the white noise at each level is randomly generated, the energies at each level and the energy for the recovered data in the time domain are shown in Table 6. The same conclusion as that in Experiment 2 can be obtained from the data shown in this table. That is, the errors at each level in the wavelet domain are not magnified when transformed back to the time domain.

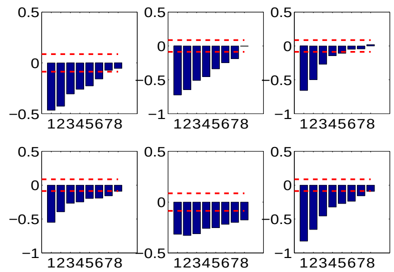
The AR processes for input data sets with different lengths are shown in Figure 39.



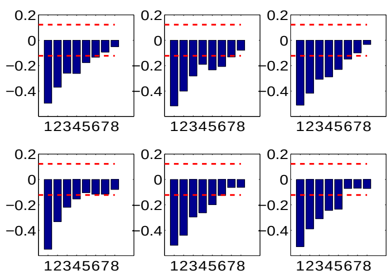
$(L = 2^{11})$



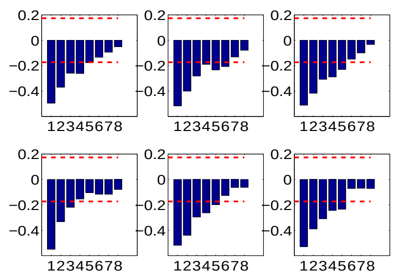
$(L = 2^{10})$



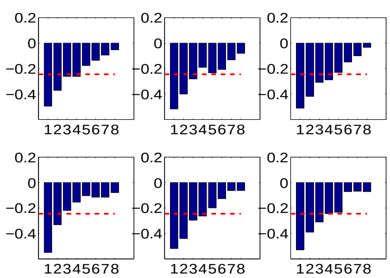
$(L = 2^{9})$



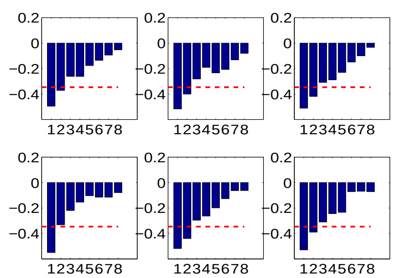
$(L = 2^{8})$



$(L = 2^{7})$



$(L = 2^{6})$



$(L = 2^{5})$

Figure 38: AR Model for Time Series of Different Lengths

Table 6: Energy at Each Level and the Recovered data

Location	Energy	Energy	Energy
1st Level	1791.7	2001.3	1249.3
2nd Level	2121.0	294.5	640.6
3rd Level	1492.8	755.7	496.7
4th Level	481.0	1783.0	274.0
5th Level	777.3	249.2	1637.3
6th Level	1983.8	417.0	901.8
7th Level	2085.9	515.2	2115.1
8th Level	0.0	0.0	0.0
Recd Time Series	631.2259	158.9566	205.1494

Thus, the conclusion obtained from experiment 4, that there exists a log-linear relationship between the length of input data and the order of the resulting AR process, is still valid for white noise with random variance.

From the above experiments, the following conclusions can be made:

- Modeling errors can be accurately estimated by an AR process,
- The order of the AR process is log-linear with the length of the input data set.
- The AR process is very robust with respect to the type of wavelet filter used in the transform.
- The errors are not magnified during the WT and inverse WT processes, that is, the errors in the time domain have a magnitude comparable to those at each level in the wavelet domain.

This provides a way to derive a systematic component by the AR process from the modeling error and feed this AR process to the original forecasting model to improve the accuracy of the forecasting results.

3.3.3 Block Bootstrapping Estimate of the LCC

The forecasting methodology provides the forecasting data for the DM process. Based on the information provided by the forecasting data, the optimal SOS and SMS are selected.

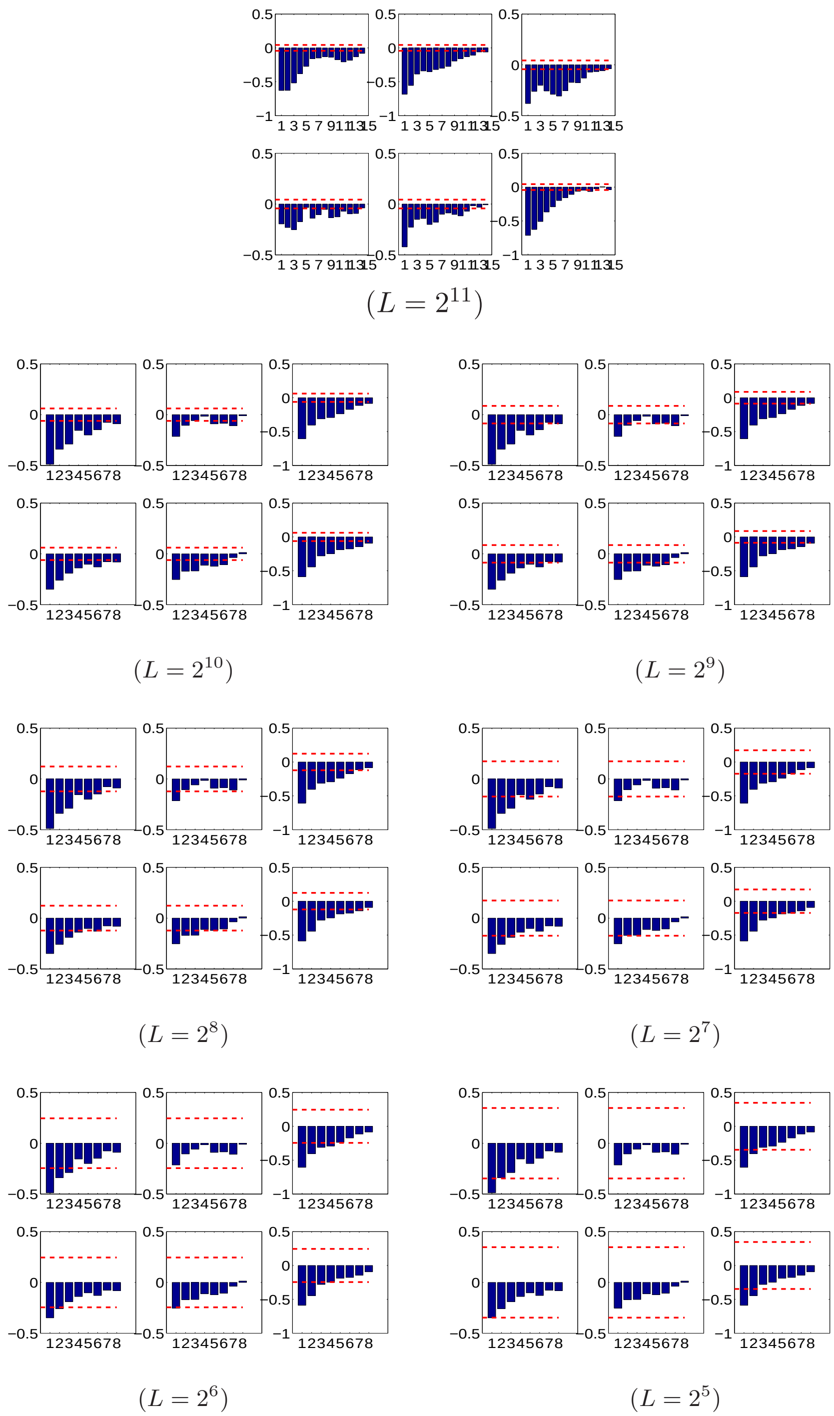


Figure 39: AR Model for Time Series of Different Lengths with Randomly Generated Variance

An estimate of the LCC of driving the business for a power plant over the planning time horizon is provided. This is a point estimate of the LCC that the power plant will actually spend in the future. The bias of the estimated LCC to the actual value, which measures the over/underestimates of the actual LCC on average, should be evaluated. Bootstrapping is a nonparametric resampling method for statistical inference commonly used to estimate confidence intervals, but it can also be used to estimate the bias and the variance of an estimator or calibrate hypothesis tests. Bootstrapping is widely used in two cases: when the use of analytical treatment is impossible, and when the data comes from a single run or limited number of runs. Papers that illustrate the diversity of recent environmentric applications of the bootstrap can be found in toxicology [8], fisheries surveys [89], groundwater and air pollution modeling [5] and [26], chemometrics [109].

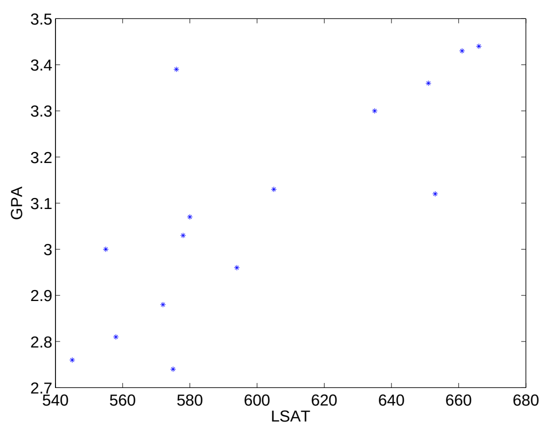
Nonparametric time series methods such as bootstrapping are becoming increasingly popular as they retain the empirical structure of the observed variables. Nonparametric methods differ significantly from the parametric alternatives because the parametric methods require assumptions regarding

- The marginal probability distribution of the variables.
- The spatial and temporal covariance of structure of the variables.

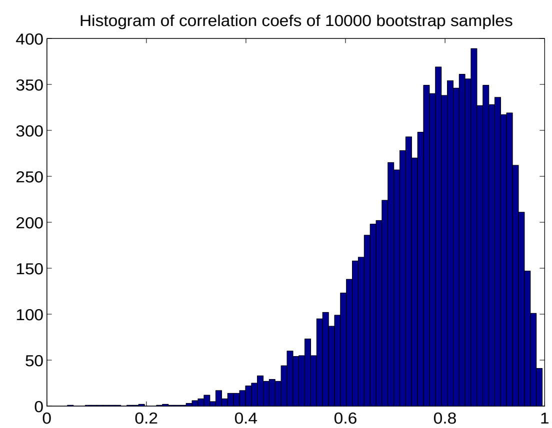
More importantly, parametric methods require estimates of various model parameters that nonparametric methods either minimize or avoid altogether. Errors arising from parameter estimation of time series models can easily overwhelm the issues of model choice [93], [94], and [105].

The methods available for implementing the bootstrap and the accuracy achieved relative to first-order asymptotic approximations depend on whether the data are a random sample from a distribution or a time series. Bootstrapping can be implemented by sampling the data randomly with replacement or by sampling a parametric model of the distribution of the data when handling data from a random sample. The distribution of a statistic is estimated by its empirical distribution under sampling from the data or parametric model. [9], [42], [30], and [28] provide detailed discussions of bootstrap methods and their properties for data that are sampled randomly from a distribution.

One example [30] that compares Law School Admission Test (LSAT) scores and subsequent law school grade point averages (GPA) from a sample of 15 law schools. The left graph of Figure 40 shows the data. The least squares fit line indicates that higher LSAT scores go with higher law school GPAs. However, how certain is this conclusion? The plot provides some intuition, but nothing quantitative. The correlation coefficient of the variables can be calculated to be 0.7764, describing the positive correlation between LSAT and GPA. Although 0.7764 may seem large, its statistically significance is still unknown. Bootstrapping can be used to resample the LSAT and GPA vectors as many times as desired so that the variation in the resulting correlation coefficients can be observed. A histogram of the result is shown in the right graph of Figure 40. Nearly all the estimates lie on the interval [0.4 1.0], providing strong quantitative evidence that LSAT score and subsequent GPAs are positively correlated. Moreover, this evidence does not require any strong assumptions about the probability distribution of the correlation coefficient.



(GPA vs. LSAT)



(Histogram)

Figure 40: Data and Histogram

The situation becomes more complicated when the data represent a time series because bootstrap sampling must be carried out in a way that suitably captures the dependence structure of the time series. The challenge is how to resample the data so that the temporal and spatial covariance structure of the original time series can be preserved. When no priori knowledge about the data is available, the best way of dealing with dependencies is the block bootstrap. This method is used to preserve the original time series structure within a block. See [54] for more information.

The basic idea of the block bootstrap is to partition the time series into blocks of observations and sample the blocks randomly with replacement. Because the sampling is done with replacement, some blocks in the data series are selected two or more times and others are not selected at all. When this process is repeated a hundred times or more, pseudo-samples that behave similarly to the underlying distribution of the original time series can be obtained. These pseudo-samples can be used in the following ways [38]:

1. Estimate the mean of these pseudo-samples, which should be close to the estimate itself.

2. Estimate the standard deviation of these pseudo-samples, which gives a bootstrap standard error of the estimate. This standard error does not rely on any distributional assumption (e.g., normality).

3. Compute the 2.5 percentile and the 97.5 percentile of these pseudo-samples, which produces a bootstrap confidence interval. The classic formula for the confidence interval can be used.

Implementation of block bootstrapping for data with an dependent structure typically requires the selection of a block length or an expected block length λ. See the related work [53], [58], [75], [76], and [18]. In recent years, various block bootstrap methods that attempt to reproduce different aspects of the dependence structure of the observed data in the pseudo samples have been proposed.

1. The moving block bootstrap; see [53], and [58].

2. The non-overlapping block bootstrap; see [18].

3. The circular block bootstrap; see [74].

4. The stationary bootstrap (SB); see [76].

The first three resample blocks of the time series with a non-random block length. The last, SB, differs from the rest in that it uses a random block length and hence, has a slightly more complicated structure. It is known that for a given block length (expected block length, for SB), all four methods have the same amount of bias asymptotically. SB is used in this study to estimate the bias.

The SB proposed by Politis and Romano [76] is simple to apply to a univariate time series, which is the case in this study. The SB replicates the time series by concatenating blocks of observations from the original time series. The blocks are selected randomly from the original time series and have a random length with a geometric distribution. To ensure the stationarity of the bootstrap time series, whenever a block exceeds the end of the time series, one continues by adding observations starting from the beginning of the time series.

Let $\mathbf{X_N} = \{X_i : i = 1, 2, \dots, N\}$ be the available observations from the sequence $\{X_i : -\infty < i < \infty\}$ with $E[X_1] = \mu$, where $X_i \in \mathbb{R}^d$ for each integer i and some integer d satisfying $1 \leq d < \infty$. Suppose that $\hat{\theta}_N$ is an estimator of the parameter of interest θ.

The block length $\lambda_i : 1 < \lambda < N$ is sampled from the geometric distribution with parameter $p \in (0, 1)$. Let $\mathbf{P}_c$ and $\mathbf{E}_c$ be the conditional probability and the conditional expectation. $\mathbf{P_c}(\lambda_1 = k) = (1-p)^{k-1}p$ for $k = 1, 2, \dots$. Also, let $I_1, \dots, I_N$ be conditionally independently and identically distributed (i.i.d) random variables with the discrete uniform distribution on $\{1, \dots, N\}$. Given the observations $\mathbf{X_N}$, the time series $\{X_i^s\}_{i\geq 1}$ is formed by periodic extension, where $X_i^s = X_j$ if $i = mN + j$ for some integers $m \geq 0$ and $1 \leq j \leq N$. Also define the blocks of length $k \geq 1$ based on the time series $X_1^s, X_2^s, \dots$ by $\mathcal{B}(i, k) = (X_i^s, \dots X_{i+k-1}^s), \quad i \geq 1, \quad k \geq 1$. Then the SB resamples $K \equiv inf\{k \geq 1, \lambda_1 + \cdots + \lambda_k \geq n\}$ blocks, given by

$$\mathcal{B}(I_1, \lambda_1), \dots, \mathcal{B}(I_K, \lambda_K).$$

Since $\mathbf{E}_c[\lambda_1] = 1/p = l$ under the geometric distribution of λ_1, on average, the lengths of the resampled blocks tend to go to infinity with N. The first N elements in the array $\mathcal{B}(I_1, \lambda_1), \dots, \mathcal{B}(I_K, \lambda_K)$ yield the SB sample $X_1^*, \dots, X_N^*$.

Define $T_N \equiv \hat{\theta}_N - \theta$, $\overline{X}_N = N^{-1} \sum_{i=1}^{N} X_i$, and $\overline{X}^*_{N,p} = N^{-1} \sum_{i=1}^{N} X^*_i$, then

$$E_*(\overline{X}^*_{N,p}) = \overline{X}_N \tag{42}$$

$$T^*_{N,p} = H(\overline{X}^*_{N,p}) - H(\overline{X}_N) \tag{43}$$

where $H : R^d \rightarrow R$ is the smooth function, so that $\theta = H(\mu)$ and $\hat{\theta}_N = H(\overline{X}_N)$. Then the bootstrap estimator of bias $(\hat{\theta}_N)$, based on the SB method described above, is given by

$$Bias(pl) = E_* T^*_{N,p}. \tag{44}$$

This process can be briefly illustrated in Figure 41.

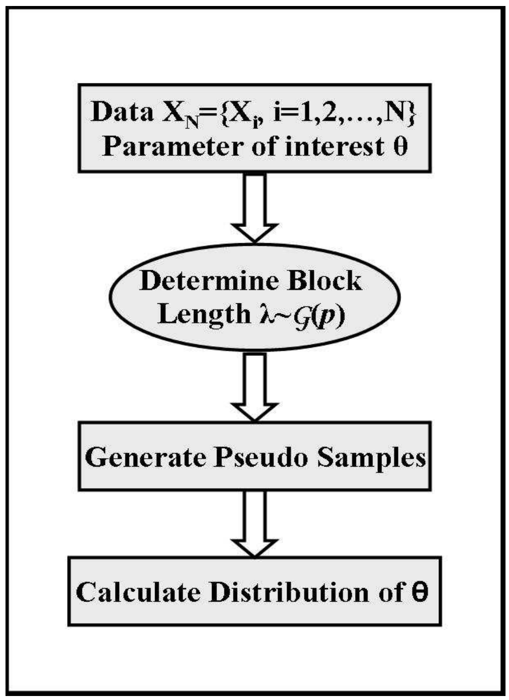


Figure 41: Block Bootstrap Process

3.4 *Uncertainty Exploration*

The electric power plant planning problem is formulated in the usual way, where the total sum of the investment, operating, and maintenance costs over a planning period is minimized through identifying the optimal SOS, SMS, and SCEP. Static formulations (deterministic) are considered first. The solution of the deterministic model produces the optimal SOS with a restriction that the system capacity must be greater than peak demand in the power plant. The optimal number of required new generation units for future years is based on customer demand forecasts.

As the electric market moves from a government-regulated monopoly to a competitive free enterprise industry, traditional planning tools that guide decision makers in developing SMS, SOS, and SCEP are inadequate because these methods often disregard the uncertainty surrounding the market environment. Fuel resource requirements, electricity prices, and customer demand are the critical input variables to the whole problem, but they are stochastic in nature. When comparing present values with forecasts made some years ago, huge deviations are evident. Forecasts for other variables also include great uncertainty from economic factors, such as interest and inflation rates, to technical considerations, such as the availability and costs of alternative generation and new emission reduction technologies. These uncertainties impose an additional risk on long-term planning because of the large influence of operational decisions. Thus, uncertainty is being more strongly considered in addressing power plant planning problems.

Different methods of handling these uncertainties have been developed according to the models analyzed. Some planning models are deterministic, using fixed values for parameters determined by more or less complex estimations. However, the estimations are usually proven erroneous. The deviations are usually large and skewed toward the optimistic side. Hence, the normal way of incorporating uncertainty into these models, sensitivity analysis, becomes invalid since it considers small ranges of variation of these parameters, and therefore, it cannot detect larger variations in some cases [57]. The other typical way of incorporating uncertainty has been by probabilistic analysis, but it is usually more complex from a computational point of view. In reality, it is usually difficult or sometimes impossible to assign probabilities to each of the different situations considered.

As a result of such difficulties, the use of scenarios has become more and more recommended, especially in rapidly changing environments such as the electric power industry. Scenario analysis was originally developed for strategic military purposes [80]. In the words of [47], a scenario is a "hypothetical sequence of events constructed for the purpose of focusing attention on causal processes and decision-points"; it considers a scenario to be a descriptive narrative of a set of relevant factors that describe alternative representations of future socio-economic conditions from a probabilistic point of view. The analysis of the scenarios help decision makers understand the role of uncertainty, explore alternative futures, and therefore, make more informed decisions in uncertain contexts. Thus, scenario analysis overcomes the DM problems by acknowledging the uncertain business environment. The emphasis of scenario analysis is not on obtaining "correct" solutions, but on designing strategies that may respond efficiently to possible changes.

Scenario analysis provides a structure for new data, frames uncertainty and balances the known with the unknown. Scenario analysis accomplishes the following:

- Creates a structure in events in the environment
- Identifies uncertainty
- Creates a structure of diverse view points
- Takes into account available knowledge
- Combines external perspectives

Generally, scenario analysis benefits the following:

- Long-term development: more robust organizational system withstanding better unexpected shocks.
- Short-term development: increased adaptability by more skillful observations of the environment.

Scenario analysis is particularly useful for analyzing the current electric business, characterized by a significant level of uncertainty regarding critical market forces and deregulation. It can provide a structured framework for imagining and assessing uncertainty, which allows the distillation of complex market interactions into a limited number of plausible alternatives that can be used to determine most appropriate strategic initiatives. By being alert to the trigger points that might signal the rise of a specific scenario, decision makers can increase their preparedness for changes in the market. More specifically, scenario analysis is utilized in systematic strategic thinking and planning for the complex and fast-growing electric power industry to identify market relevant factors, examine the interactions of current trends and uncertainties, and then determine a suitable strategy for this forecasting problem within a given market domain and time frame.

For purposes of scenario analysis, decision makers should select, from among the identified market forces, several forces that are anticipated to have a great potential impact on decision making. These forces should be used to form a matrix that presents all plausible futures or scenarios. Each of the scenarios examines a different possible development for the electric power plant resulting from the interactions of critical forces. By analyzing the implications of each scenario, decision makers will be able to identify or develop strategies that would be successful under various future conditions and particularly valuable in a specific scenario. By being alert to the triggers that might indicate the onset of a particular scenario, decision makers can begin to adjust their strategies to prepare for shifts in electric power plants. Through scenario analysis, major uncertainties faced by electric power plants can be addressed systematically, and a set of robust and adaptable strategies that allow the power plants to stay one step ahead of the market can be developed.

As mentioned above, the basic aim of scenario analysis is not to forecast the future or fully characterize its uncertainty, but rather to bound this uncertainty. In this sense, scenario analysis may be complemented with traditional forecasting and simulation techniques in order to provide a composite picture of future developments for use as the background for decision making or strategic planning [87]. Thus, scenario analysis, combined with the proposed forecasting method WAW, is more suitable for describing future states of the highly complex, innovative, and fast-growing electricity business.

A more structured method useful for scenario analysis is morphological analysis. Morphological analysis is a non-quantified modeling method for structuring and analyzing technological, organizational, and social problem complexes. It can be carried out in two phases. The first phase, the analysis phase, relies on the representation of a problem using a number of parameters (or variables) that are allowed to assume a number of conditions (or states). In the second phase, the synthesis phase, consistent alternatives are derived by considering the consistency between conditions for different parameters in a pair-wise fashion.

3.4.1 External Factors Identification

The first step of scenario analysis is to specify the scope and time frame of the problem. Power plants have to deal with uncertainties in all aspects of system operation and planning, especially in long-term system planning, due to the large risk involved in the DM process. A deterministic DM process that describes the evolution of the electric power plant under "normal conditions" is formulated. The basis of this decision is the future conditions of such input forecasted without considering the impact of the variations in the external environment. Nevertheless, uncertainty in the external environment will have an undeniable impact on forecasting results and consequently on the DM process, necessitating the identification of the external factors, including social, economic, environmental, political, and technological factors that are most relevant to DM process. The SEEPT framework is an efficient way to obtain a holistic view of the many forces that will affect a single system such as a power plant. Based on the effects of such forces on the evolution of the system, they can be categorized into the following two groups:

- Specific events, such as, the passage of legislation.
- General trends, such as, an increase in the cost of fuel.

The list of external forces can yield as many as 50 driving forces. The next step is to analyze and prioritize these forces based on their level of predictability and importance in affecting the desired outcome. This step reduces this large set of forces to only those most relevant to the decision focus. A logical and rigorous thinking through of the forces and trends often helps identify the forces that are most relevant to the decision without complex analyses.

In summary, the tasks that must be accomplished during this step include the following:

- Specify the scope of the planning and the time frame.
- For the present situation, develop a clear understanding that will serve as the baseline for each of the scenarios.

- Identify predetermined elements that are virtually certain to occur and that will be driving forces.
- Identify the critical uncertainties in the environmental variables.
- Identify the most important drivers.

3.4.2 Scenarios Generation

After the identification of driving factors, the next step is to identify the possible conditions for each. A morphological field that represents the problem, its parameters, and conditions is then utilized. The parameters are shown in the columns, with boxes representing representing possible conditions (see Figure 42). A given alternative, in which conditions are assigned to each parameter, is shown by highlighting the relevant condition for each of the parameters. The alternative is characterized by $\{X3, Y4, Z1\}$, representing only one of $5 * 5 * 3 = 75$ possible scenarios for this morphological field. One of these scenarios most likely will reflect the mainstream views of the future. The other scenarios will shed light on other possibilities.

Parameter X	Parameter Y	Parameter Z
X1	Y1	Z1
X2	Y2	Z2
X3	Y3	Z3
X4	Y4	
X5	Y5	

Figure 42: Morphological Field

Representing the condition of a parameter in one dimension is normally problematic. For example, the influences of weather and economic factors, two important external driving forces in the electric business, depend not only on their values but also on the time in which they occur and how long their effects last. Thus, each condition of the weather is expressed as a vector $W = [v, t, d]$, where v represents the value of the weather, temperature, t

represents when this phenomenon occurs, and d represents how long this phenomenon lasts. The ranges of each element of each parameter should be identified, so one morphological field is required for each parameter to determine its condition. Figure 43 shows the morphological fields for two parameters W and E with each property element at two levels, each parameter with 8 conditions.

Parameter W	Element v_w	Element t_w	Element d_w
	v_{w1}	t_{w1}	d_{w1}
	v_{w2}	t_{w2}	d_{w2}

Parameter E	Element v_e	Element t_e	Element d_e
	v_{e1}	t_{e1}	d_{e1}
	v_{e2}	t_{e2}	d_{e2}

Figure 43: Morphological Fields for Parameters

In this case, more than one morphological field is used to analyze the problem. The two fields can then be combined at a later stage. Two-field morphological analysis is required when the number of relevant parameters is large or when the complex problem consists of two or more separate contexts such as an external scenario field and an internal strategy field. In this study, combining these two morphological fields generate 64 scenarios. This provides a complete description of the picture with two parameters each of which has three property elements at two levels.

The synthesis phase of morphological analysis allows the elimination of a large number of scenarios by judging the consistency between the conditions for different parameters. Each scenario, which should be chosen in a systematic way, has to be internally consistent and plausible, and together, they have to cover a reasonable variety of different developments, "to span the problem space." No golden rule unequivocally gives the number of scenarios needed. The complexity of the problem, the resources available to analyze the consequences of the scenarios later in the process, and the level of detail desired in each scenario, will all affect this number.

Each scenario is used as an input to the DM process, and decisions are made based on the information provided by the scenario. Finally, the scenario analysis results are used by the decision makers for further discussion or for the solution of the problem.

In summary, the tasks that need to be finished in this step are as follows:

- Define the conditions for each key external force.
- Create morphological fields with the key forces, eliminate the inconsistent scenarios.
- List surviving scenarios that will assist in making decisions.
- Carry out the DM process with each scenario as the input information.
- Obtain results for each scenario.

3.4.3 Scenarios Analysis

This step analyzes the results of the DM process for each scenario. Through analysis, a multitude of questions must be answered:

- What is the best strategy for dealing with this situation?
- What are the major opportunities and risks in this scenario?
- What should the system do or not do when a specific scenario will take place?

By answering these questions, a series of simple contingency plans for each potential future can be developed. The next phase is to assess how much these strategies have in common with the current strategy to identify:

- Which strategic alternatives seem to be suggested by a majority or all the scenarios? These should be key parts of any strategic plan.
- Which strategic alternatives challenge most strongly the assumptions underlying the current strategy? The scenarios from which they are drawn need further consideration and provide a guidance in rethinking strategic orientation. Even if the scenarios do not come to pass, they highlight a blind spot in the current plans. When final strategic decisions are made, creative ways that include these insights should be found.

- Which strategic alternatives are logical extension to the current strategy? These strategies give decision-makers an idea of how alternate future developments could be leveraged to push forward an agenda or program that is already in place.

The final step is to decide which strategic alternatives should be adopted. Again, a simple set of questions can serve as a guide:

- What events would trigger each strategic alternative? What impact (positive or negative) would those events have on the system? How effective is the strategy at addressing these issues?
- What is the evidence to support the assumptions underlying the strategic suggestion? What aspect of the scenario serves as the underpinning of the strategy?
- Is it feasible for the system involved to execute the strategy? What would prevent it from being able to do so?

CHAPTER IV

FORECASTING RESULTS AND ANALYSIS

The forecasting method WAW was applied to the fleet management of generation units in a power plant as a proof of the implementation of the DM process. In this chapter, WAW was first utilized to provide forecasting information for customer demand, natural gas prices, and electricity prices, the three parameters on which the decisions are based. The forecasting results were then validated with the real data to prove the accuracy level. The forecasting results were then compared with the results obtained from the traditional Holt-Winters' method, and the WAW method was proven to provide overall better performance.

4.1 Customer Demand Forecasting

4.1.1 Historical Data

Historical data for customer demand within a given market domain are obtained from July 1981 through October 2002. The data set $D = \{d_t\}_{t=1,2,\ldots,n}$ consists of $n = 256$ (2^8) monthly data points corresponding to $\{(\text{July 1981}), (\text{Aug. 1981}), \ldots, (\text{Oct. 2002})\}$. Given this set of 256 observations obtained at uniformly spaced time intervals, it is often convenient to rescale the time axis in such a way that it becomes the set of integers $\{1, 2, \ldots, n\}$. This amounts to measuring time in months with July 1981 as the 1^{st} month and October 2002 as the 256^{th} one. Figure 44 shows the historical data of customer demand. The graph shows that customer demand has an upward trend and a strong seasonal pattern. Figure 45 "zooms in" on the customer demand data in $1984 - 1985$ and $1995 - 1996$. The data reveal that customer demand peaks in August and January and troughs in May and November during each year. Figure 45 also shows an increase in customer demand from the 1980's to the 1990's.

An inspection of Figures 44 and 45 suggests the possibility of representing the data as

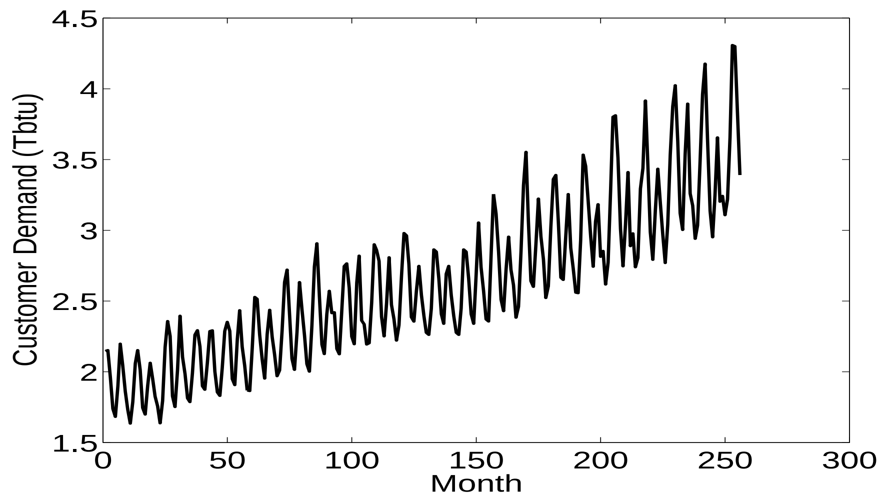


Figure 44: Residential and Commercial Demand ($Tbtu$)

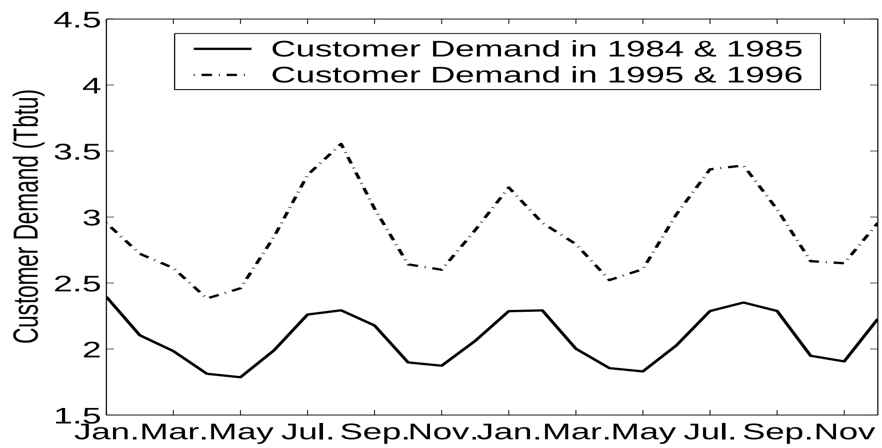


Figure 45: Seasonal Patterns Existing in the Historical Data

a realization of the process

$$d_t = m_t + s_t + h_t, \tag{45}$$

where m_t is a slowly changing function known as a trend component, s_t is a function with a known period, referred to as a seasonal component, and h_t is a stationary, high-frequency component whose mean and autocovariance function are both independent of time. However, the seasonal and high-frequency fluctuations in Figure 44 appear to increase with the process; thus, a preliminary transformation of the data is used so that the transformed data are more compatible with the classical decomposition in Equation (45). A comparison of customer demand in Figure 44 with the transformed data in Figure 46, obtained by applying a logarithmic transformation, show that the transformed data do not exhibit increasing fluctuations with increasing level, apparent in the original data. This suggests that the decomposition represented by Equation (45) is more appropriate for the transformed data than for the original series. The transformed data will be referred to as customer demand in later analysis.

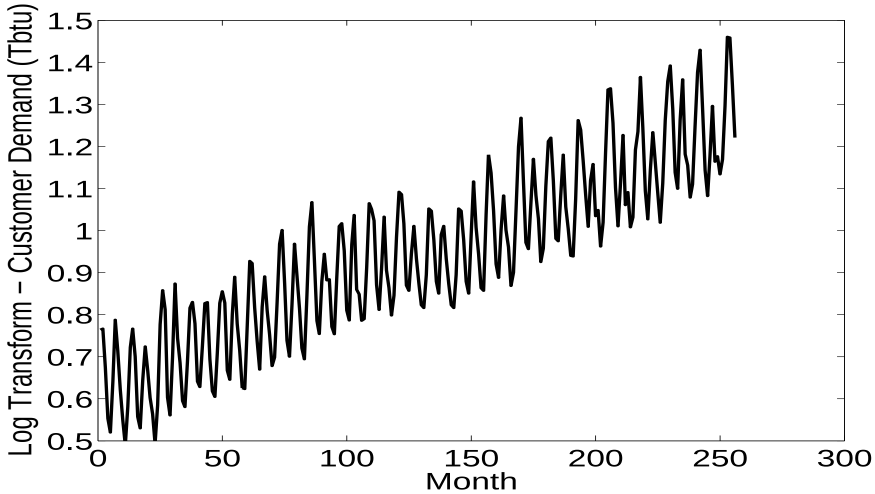


Figure 46: The Log Transform of Residential and Commercial Demand

4.1.2 Data Analysis

4.1.2.1 Wavelet Transform

The customer demand is transformed to the wavelet domain by the non-decimated wavelet transform performed with the Symmlet 8 filter. Figure 47 shows customer demand in the wavelet domain. The left graph shows the transformed data with a transform depth of five levels. The first level represents the high-frequency component existing in the data. The second and third levels obviously reflect the seasonal characteristics in the data, as illustrated by the periodic phenomena within these two levels. However, there is little seasonality left in the fourth level. The fifth level represents the trend, which is the “smooth” part of the wavelet transform. Because the fourth level reflects very little seasonality, transforming the data into four levels, not five, as illustrated in the right graph is better. The first level represents the high-frequency component in the data. The second and third levels explain the seasonality, but the fourth level represents the trend. A comparison of these two transforms shows that the four-level tranform provides enough information through decomposing the data into trend, seasonal, and high frequency components. Therefore, the latter transform is used in the customer demand data analysis.

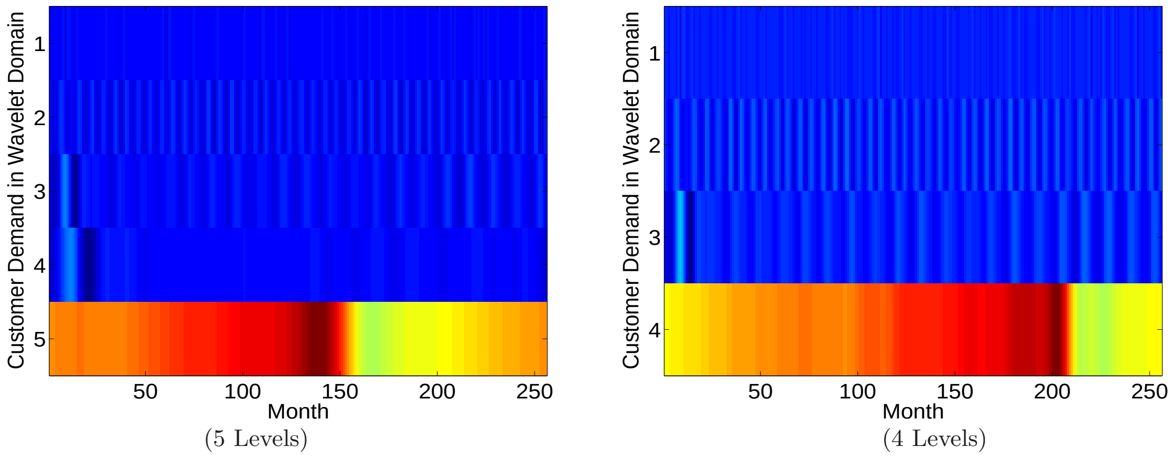


Figure 47: Customer Demand in the Wavelet Domain, Performed with Symmlet (8)

In the wavelet domain, the data are decomposed into high-frequency, seasonal, and trend components. For each component, an analysis is done and an appropriate method that simulates its behavior is found. Utilizing the information obtained through the analysis, the following 24 months of data are provided.

4.1.2.2 First Level Data Analysis

The first level is shown in the left part of Figure 48. An autoregressive process is utilized to simulate the data, but the order of the AR process generating the data is not known. It might be possible that there is no ***true*** AR process, so the goal is to find one that represents the data optimally in some sense. The right part of Figure 48 plots the sample PACF together with the bounds $\pm 1.96/\sqrt{n}$. From this graph it is easy to read off the preliminary estimator of $p = 8$.

Yule-Walker procedure is known to be applicable to the fitting of AR processes. It is used to estimate the coefficients of this AR(8) process. With equations

$$\hat{\phi} = (\hat{\phi}_1, \ldots, \hat{\phi}_p)' = \hat{R}_p \hat{\rho}_p, \text{ and} \tag{46}$$

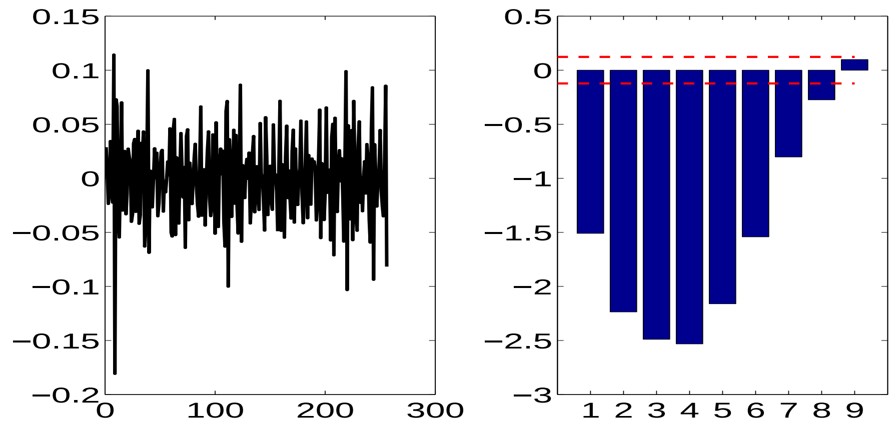


Figure 48: The First Level Data and Fitness Test

$$\sigma^2 = \hat{\gamma}(0)[1 - \hat{\rho}_p^{'} \hat{R}_p^{-1} \hat{\rho}_p], \tag{47}$$

where

$$\hat{\rho}_p = (\hat{\rho}(1), \ldots, \hat{\rho}(p))^{'} = \hat{\gamma}_p / \hat{\gamma}(0)_p,$$

the value of the coefficients in Equation (41) are estimated to be

$$\hat{\phi} = \{1.551, 2.338, 2.668, 2.772, 2.434, 1.805, 1.034, 0.4274\}.$$

Figure 49 shows that the AR(8) process is able to capture the main characteristics of the first level data. The forecasting results for the following 24 months are also shown in the figure.

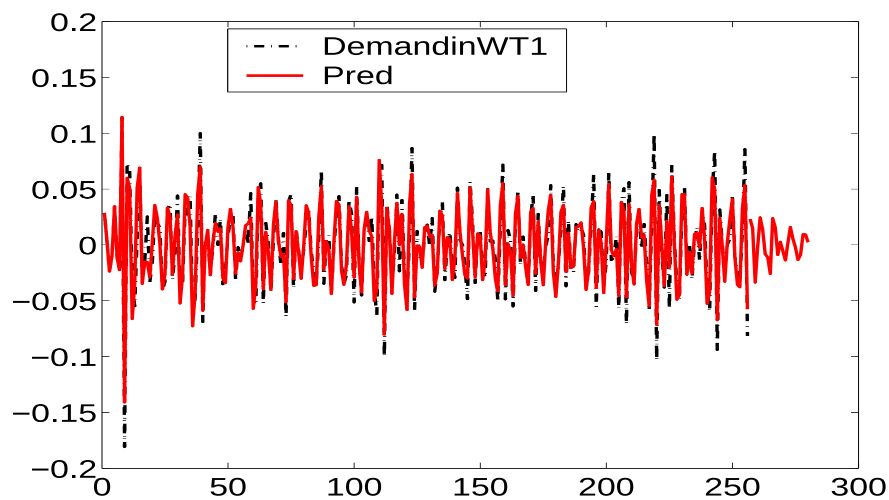


Figure 49: The First Level Predicted by the AR(8) Process

4.1.2.3 Second Level Data Analysis

The second level is part of the seasonal component in the data. Figure 50 shows a strong seasonal pattern. Harmonic regression is utilized to fit the data. The fitted model can be expressed as

$$Y = \alpha_0 + \sum_{n=1}^{N} (\beta_n \cos(n\omega x) + \gamma_n \sin(n\omega x)),$$

with $\alpha_0 = 0.000486$, and the values for β_n and γ_n shown in Table 7.

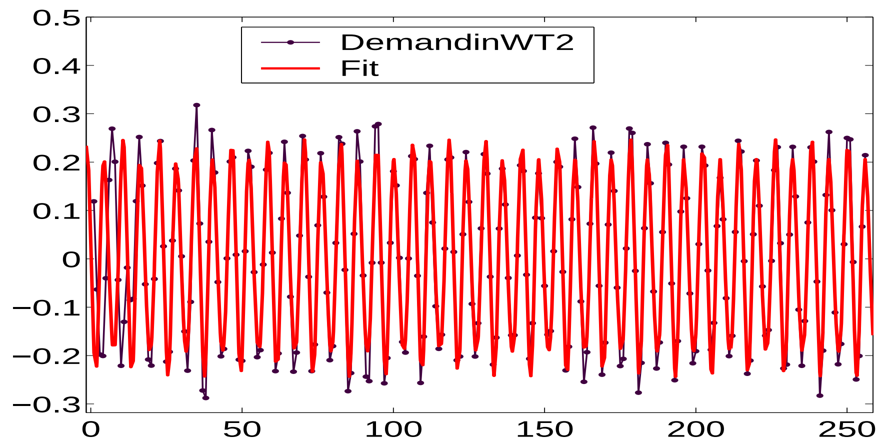


Figure 50: The Second Level Fitted Using Harmonic Regression ($\omega = 0.5244$)

Table 7: Second Level Harmonic Regression Coefficients

n	β_n	γ_n
1	0.01234	-0.004047
2	0.01203	-0.2165
3	0.01704	-0.02187
4	-0.001865	0.006126
5	-0.001689	0.002681
6	-0.002351	0.002559
7	-0.002813	0.001746
8	-0.0005374	0.00291

The frequency of the harmonic regression $\omega = 0.5244$, which implies that the period of seasonality is

$$\text{period} = \frac{2\pi}{\omega} = 11.98,$$

which is verified by the fact that natural phenomena usually occur every 12 months. Figure 50 shows the fitted model by using harmonic regression, and Figure 51 illustrates the forecasting results for the following 24 months.

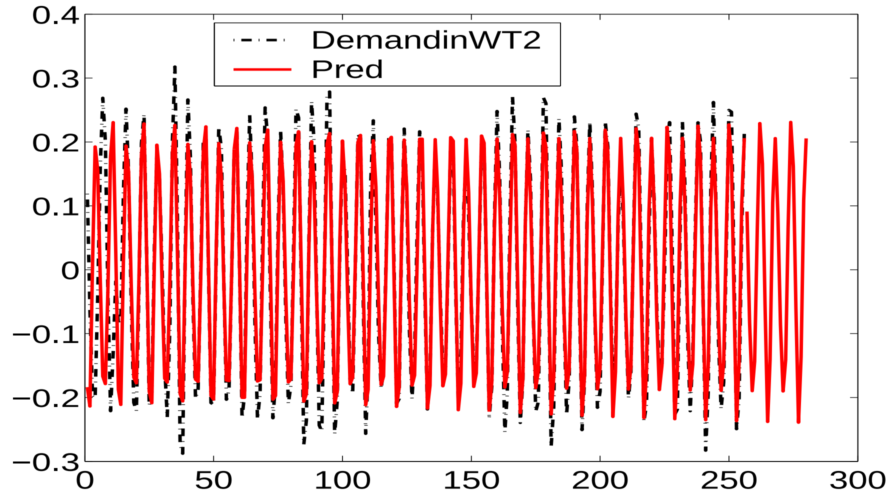


Figure 51: Second Level Forecasting Results ($\omega = 0.5244$)

4.1.2.4 Third Level Data Analysis

The third level is also fitted using harmonic regression. Figure 52 shows the original data and the fitted data using harmonic regression. The coefficients obtained are shown in Table 8.

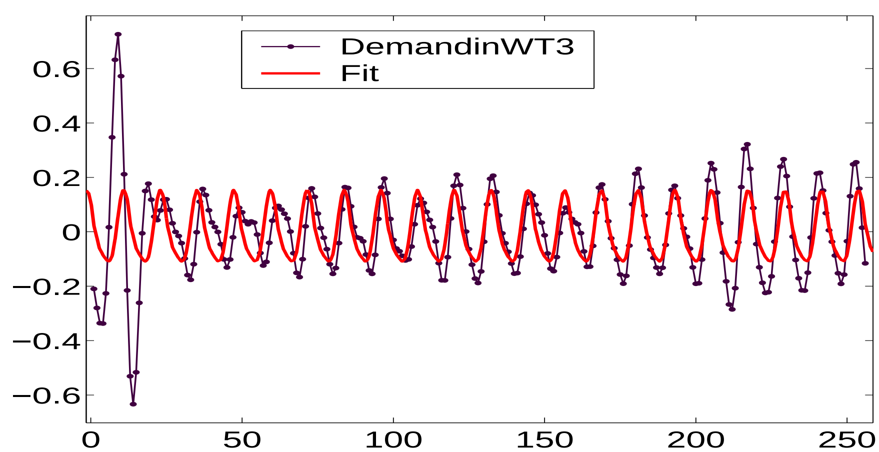


Figure 52: The Third Level Fitted Using Harmonic Regression ($\omega = 0.5174$)

The frequency of the harmonic regression $\omega = 0.5174$, which corresponds to a period of

Table 8: Third Level Harmonic Regression Coefficients

$\alpha_0 = -9.838e^{-005}$		
n	β_n	γ_n
1	0.1071	-0.06599
2	-0.0009914	-0.02927
3	0.0001788	-0.001604
4	0.0007348	-0.001198
5	0.001301	-0.0007688
6	0.001535	-0.0001318
7	0.001485	0.0005786
8	0.0008498	0.001277

12.14 months. The period for the third level is quite close to that for the second level. Both explain the seasonal characteristics in the original data. Figure 53 shows the forecasting results for the following 24 months.

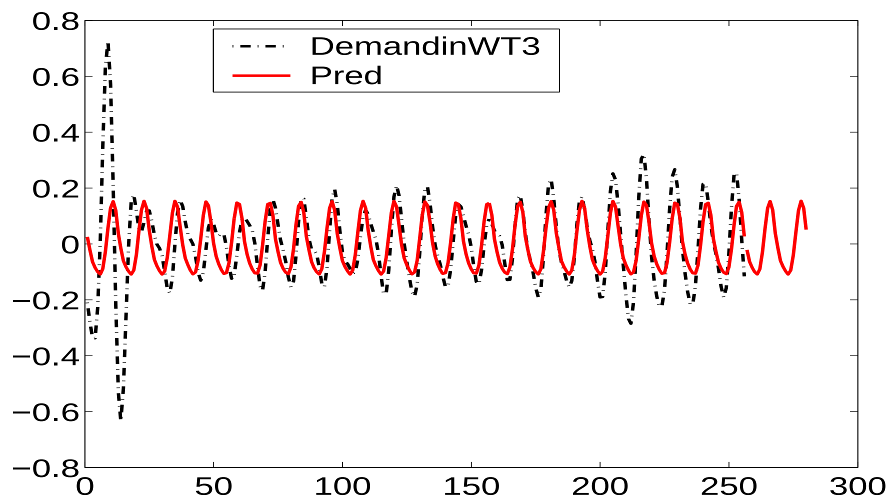


Figure 53: Third Level Forecasting Results ($\omega = 0.5174$)

4.1.2.5 Fourth Level Data Analysis

The fourth level of the data in the wavelet domain represents the trend in the original data. Holt-Winters' method is used to forecast the future. The preset parameters α, β, and γ are set to be 0.1, 0.2, and 0.3, respectively. Figure 54 shows the forecasting results of Holt-Winters' method.

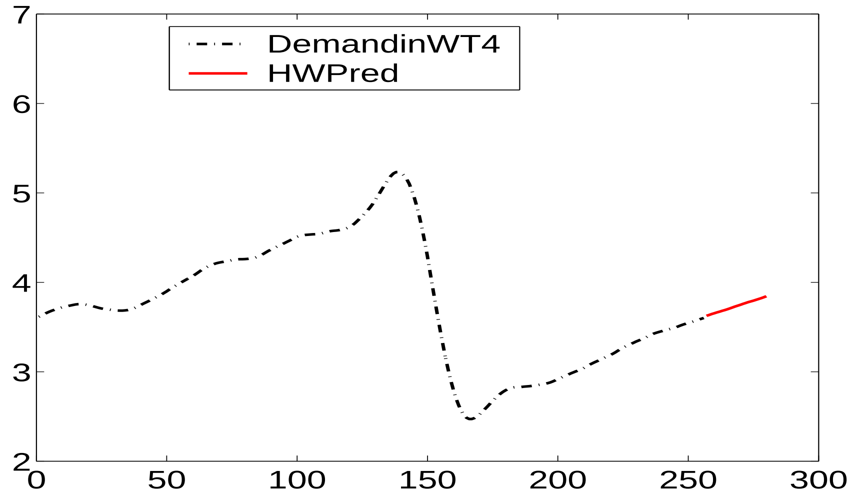


Figure 54: Fourth Level Forecasting Results Performed by Holt-Winters' Method

4.1.3 Forecasting Results

Forecasting is obtained in the time domain by combining the predicted trend, seasonality, and high-frequency component. This step involves the inverse WT of the forecasted values at different levels. Figure 55 shows the forecasting results for the following 24 months.

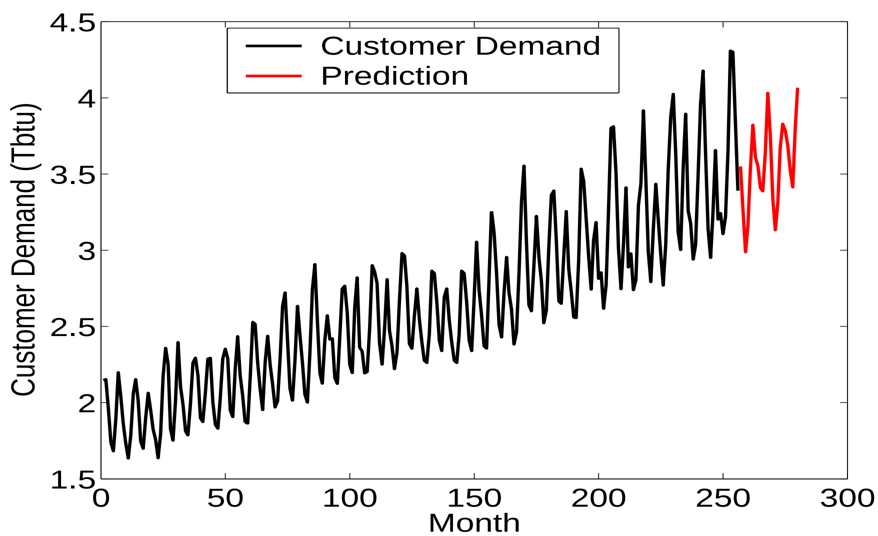


Figure 55: Forecasting Results for the Following 24 Months

4.2 Natural Gas Prices Forecasting

4.2.1 Historical Data

Historical data for natural gas prices within a given market domain are obtained from July 1981 through October 2002. The data set $FC = \{fc_t\}_{t=1,2,\ldots,n}$ consists of $n = 256$ (2^8) monthly data points corresponding to $\{(\text{July } 1981), (\text{Aug. } 1981), \ldots, (\text{Oct. } 2002)\}$. Figure 56 shows the historical data of the natural gas prices with converted time axis $\{1, 2, \ldots, 256\}$. The natural gas prices appear to fluctuate erratically about a slowly changing level. The average natural gas for this market domain was as low as \$2.71/Mcf from 1981 to 1999 whereas at the intersection of 2000 and 2001, the natural gas price soared to as high as \$9.47/Mcf. This increase reflects a competitive market reaction as supply lagged in response to a recent surge in demand. Gas demand in 2000 increased due to a number of factors, including the start of operations at new gas-fired electric-power generators and new home construction, which tends heavily toward the use of natural gas for heating and cooking. The seasonal pattern is not that apparent in the data in the time format as in those in customer demand. This pattern is partially due to the fact that natural gas is used primarily for manufacturing and electric power generation, as well as in residential cooking and water heating during the summer. But residential heating requirements increase the total demand for natural gas in excess of production and import capabilities during the winter.

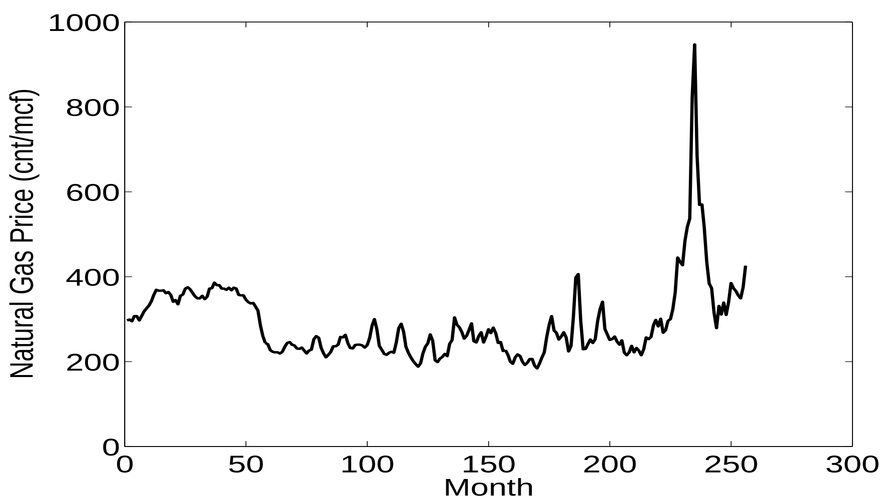


Figure 56: Natural Gas Electric Utility Purchase Prices (cnt/mcf)

4.2.2 Data Analysis

4.2.2.1 Wavelet Transform

The natural gas prices are transformed to the wavelet domain by the NDWT performed with the Symmlet 8 filter to extract critical information for forecasting. Figure 57 shows the natural gas prices in the wavelet domain. The left graph shows the transformed data with a transform depth of five levels, the right one with a transform depth of four levels. A comparison of these two transforms shows that the one with a transform depth of four levels provides enough information through decomposing the data into high-frequency (first level), the seasonal (second and third levels) and the trend (fourth level) components. Therefore, the four-level transform is used in the data analysis of natural gas prices.

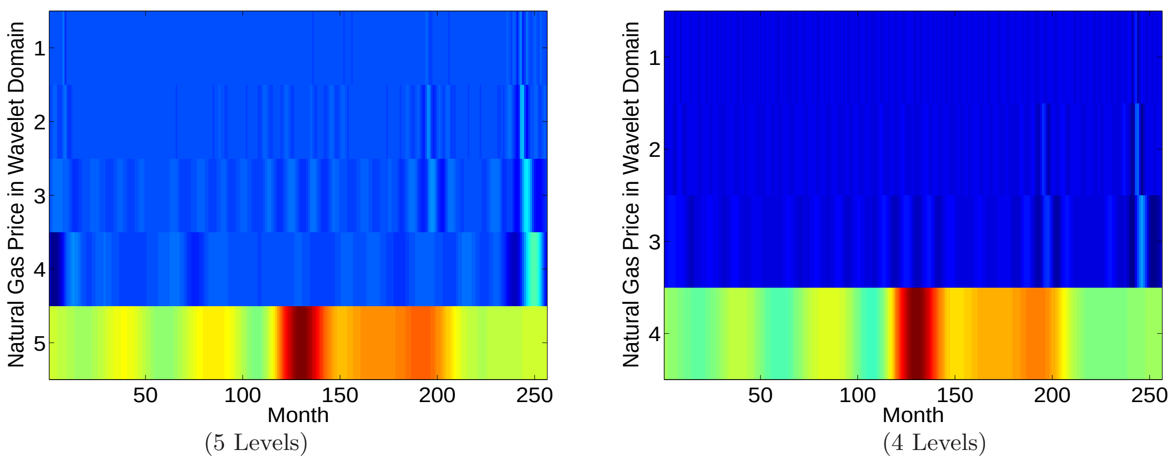


Figure 57: Natural Gas Prices in the Wavelet Domain, Performed with Symmlet (8)

4.2.2.2 First Level Data Analysis

The first level is shown in the left graph of Figure 58. The big spike in the original data is captured by the first level data. Such a phenomenon is usually hard to model by traditional methods. In the WAW method, external factors can be introduced through the use of the ARMAX model. This is done through the following steps:

- Test the order of the AR part of the ARMAX model. Yule-Walker model is applied to estimate the order of it. The right graph in Figure 58 shows the sample PACF together with the bounds $\pm 1.96/\sqrt{n}$. From this graph, it is easy to read off the preliminary estimator of $p = 7$.

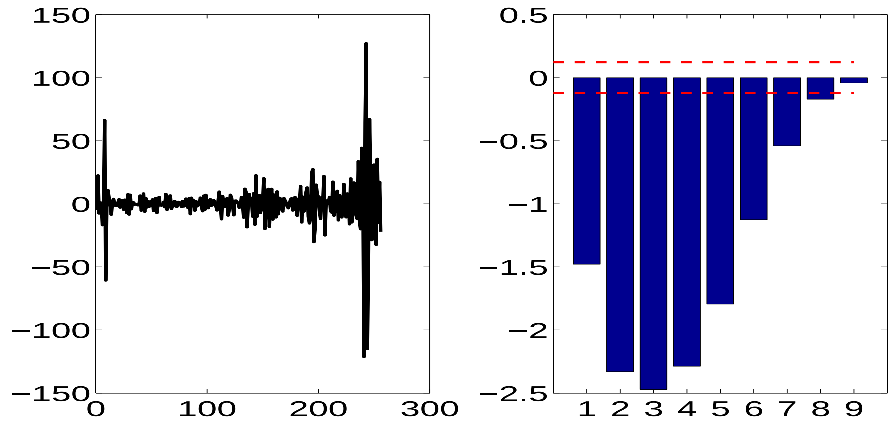


Figure 58: The First Level Data and Fitness Test

- Introduce an external factor. Figure 59 plots the the external factor in the bottom graph, which is treated as an input to the ARMAX model. The first level data are plotted in the upper graph and treated as the output of the ARMAX model.

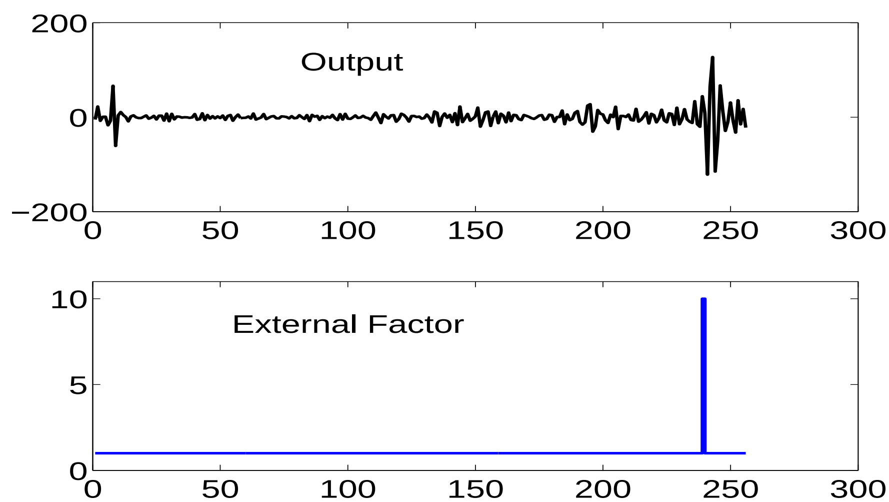


Figure 59: The First Level Data and the External Factor IDPlot

- Determine the time lag of the external factor with respect to the first level data. The sample cross-correlation function (XCF) between them is computed. XCF is a vector of length $2{*}nLags+1$ corresponding to lags $0, \pm 1, \pm 2, \ldots, \pm nLags$. The center element of the XCF contains the zeroth lag cross correlation. A two element vector called the bounds indicates the approximate upper and lower confidence bounds, which assume that the two series are completely uncorrelated. Figure 60 shows that the cross-correlation function peaks at the fourth lag. ARX[7,7,0] is used to model the first level data, which is a special case of ARMAX when the order for the moving average MA process is zero with

$$\hat{\phi} = \{1.729, 2.428, 2.597, 2.273, 1.616, 0.8973, 0.3336\}, \text{ and}$$

$$\hat{\xi} = \{2.276, 5.039, -4.907, -5.877, 3.629, 1.605, 0.05419\}.$$

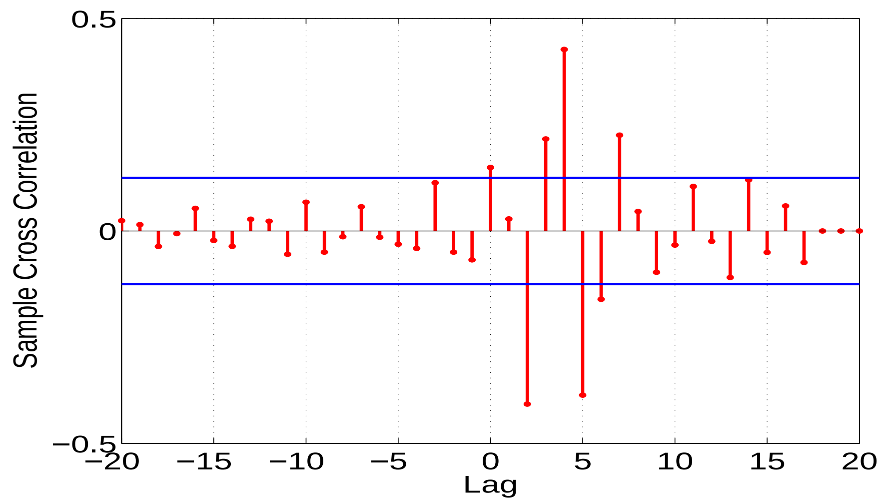


Figure 60: Correlation between the First Level Data and the External Factor

- Inspect the goodness of the model. Figure 61 shows that the model residuals are not correlated within themselves in the upper graph. The lower graph shows that the residuals are not correlated with the external factor. These are implied by the small amplitude of the correlation functions, which is a good model feature.

The ARX model developed through the above procedures is utilized to perform the forecasting. The forecasting results are shown in Figure 72. The figure shows that the developed ARX model can well simulate the big spike in the historical data by introducing an external factor.

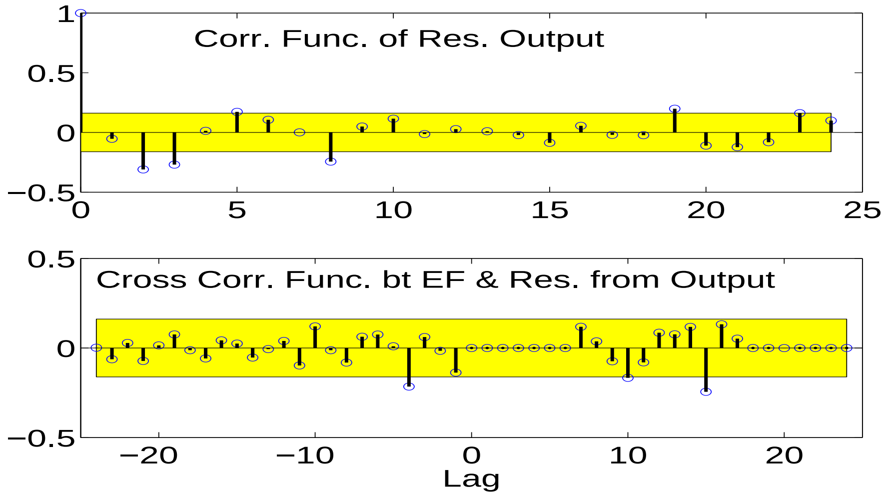


Figure 61: Correlation Relationship of the Residuals

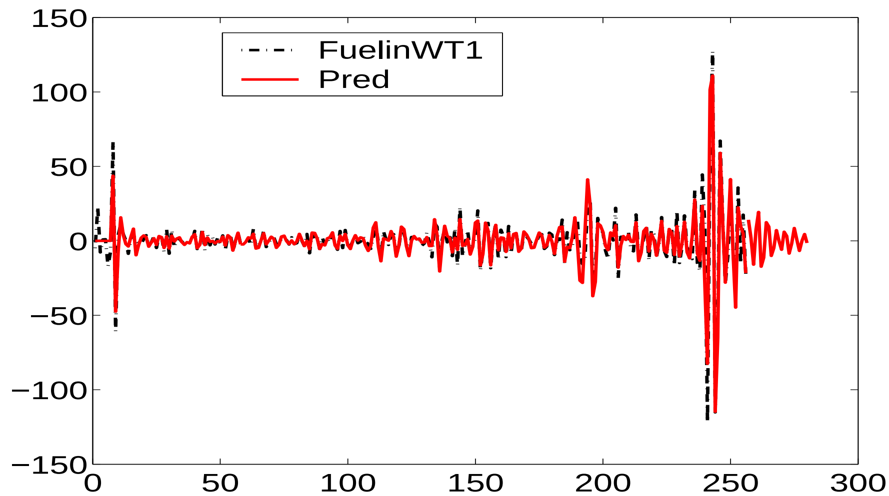


Figure 62: The First Level Data Fitted Using ARMAX Process

4.2.2.3 Second Level Data Analysis

The second level captures the seasonality in the historical data. Figure 63 shows the second level data with obvious seasonal variations, but the big spike in the original data is also captured. Harmonic regression is used to account for most of the cycles present in the data. Table 53 calculates the coefficients for harmonic regression, with a period of 12 months. The problem with simple harmonic regression, however, is that the coefficients are fixed with time, shown in Figure 63. It cannot simulate the variations in the amplitude. Therefore, depending solely on harmonic regression can not provide satisfactory results.

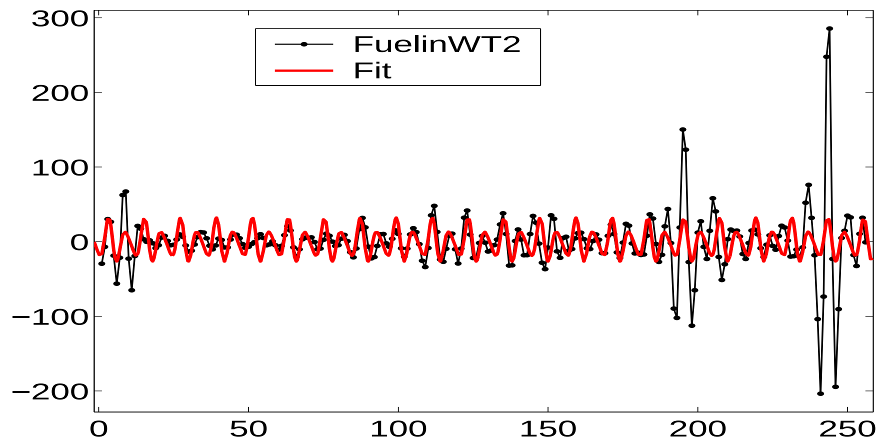


Figure 63: The Second Level Data Fitted Using Harmonic Regression ($\omega = 0.5233$)

Table 9: Second Level Harmonic Regression Coefficients

$\omega = 0.5233$		
$\alpha_0 = -0.1616$		
n	β_n	γ_n
1	-0.2606	4.198
2	-21.62	-3.208
3	5.262	-4.273
4	0.191	0.1925

Gaussian regression, used to capture the envelope of the second level data, is used for fitting peaks, given by

$$Y = \sum_{n=1}^{N} a_n e^{-(\frac{x-b_n}{c_n})^2},$$

where $\{a_n\}_{n=1}^N$ are the amplitudes, $\{b_n\}_{n=1}^N$ are the centroids (or locations) of each peak, $\{c_n\}_{n=1}^N$ are related to the peak width, N is the number of peaks to fit, and $1 \leq N \leq 8$. Figure 64 and Table 10 show the results of the upper envelope, with 3 peaks to fit. Figure 65 and Table 11 show the results of the fitting of the lower envelope, also with 3 peaks to fit.

Table 10: Second Level Upper Envelop Gaussian Regression Coefficients

n	a_n	b_n	c_n
1	276.1	242.7	3.927
2	75.77	195.4	9.636
3	65.57	8.021	4.848

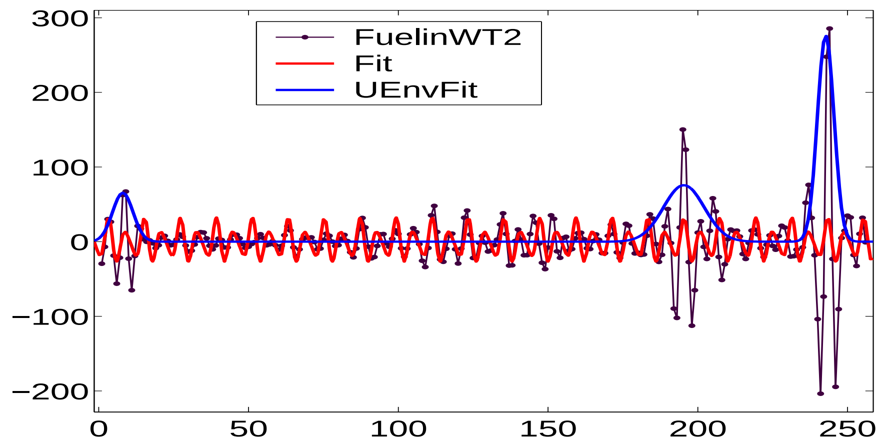


Figure 64: Second Level Upper Envelop Fitted Using Gaussian Regression

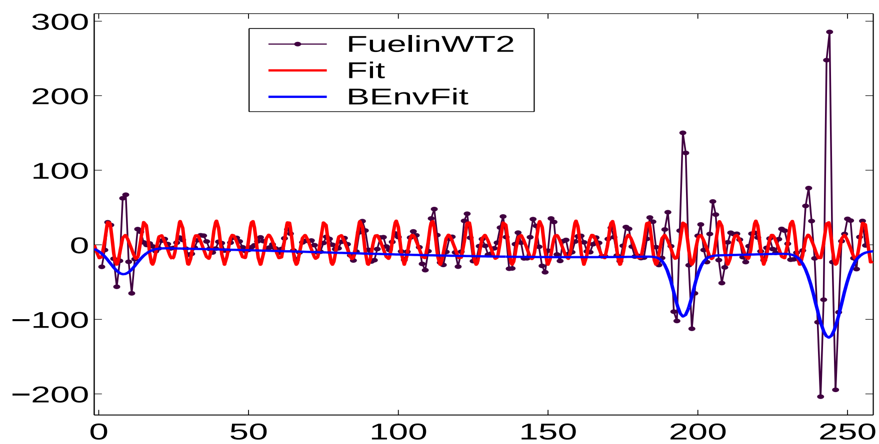


Figure 65: Second Level Bottom Envelop Fitted Using Gaussian Regression

Table 11: Second Level Bottom Envelop Gaussian Regression Coefficients

n	a_n	b_n	c_n
1	-114.3	243.7	6.054
2	-81.14	195.2	4.436
3	-36.04	8.241	6.119

Combining harmonic regression with Gaussian regression provides results that capture not only the seasonal variations but also amplitude variations. Figure 66 shows the forecasting results from this process.

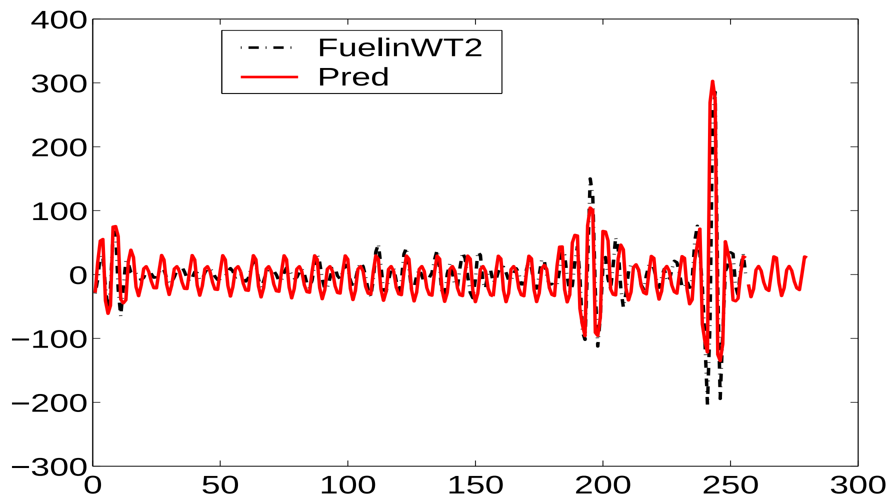


Figure 66: Second Level Forecasting Results ($\omega = 0.5233$)

4.2.2.4 Third Level Data Analysis

The third level also represents seasonal characteristics. The period of seasonal variation is calculated to be 12.08 months through harmonic regression, which is consistent with that calculated in the second level. Gaussian regression is utilized to simulate the amplitude variations. Figures 67 to 69 show the regression results. Tables 12 to 14 present the corresponding regression coefficients.

Table 12: Third Level Harmonic Regression Coefficients

$\omega = 0.520$		
$\alpha_0 = -0.08649$		
n	β_n	γ_n
1	-28.06	42.23
2	-3.148	-3.63

Table 13: Third Level Upper Envelop Gaussian Regression Coefficients

n	a_n	b_n	c_n
1	304.8	244	11.04
2	107.1	197.9	12.51
3	57.24	127	41.84
4	729.3	-191.9	125

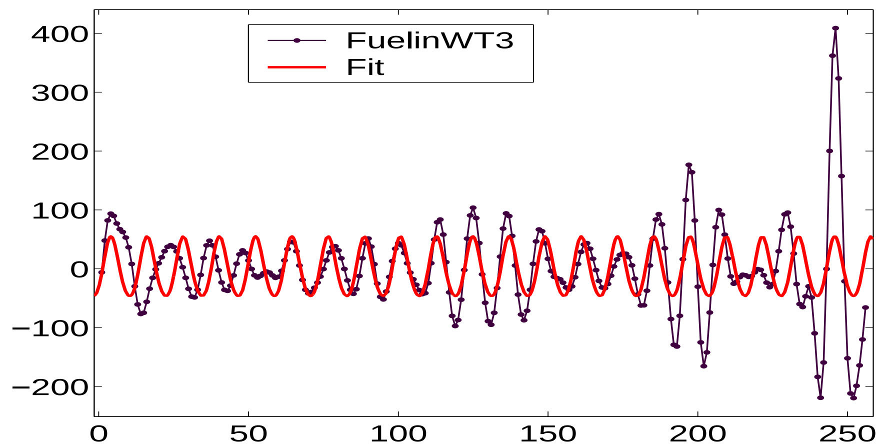


Figure 67: The Third Level Fitted Using Harmonic Regression ($\omega = 0.520$)

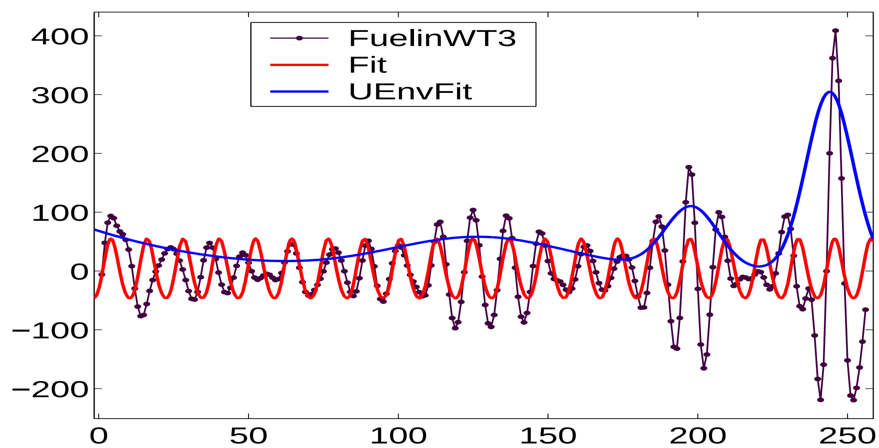


Figure 68: Third Level Upper Envelop Fitted Using Gaussian Regression

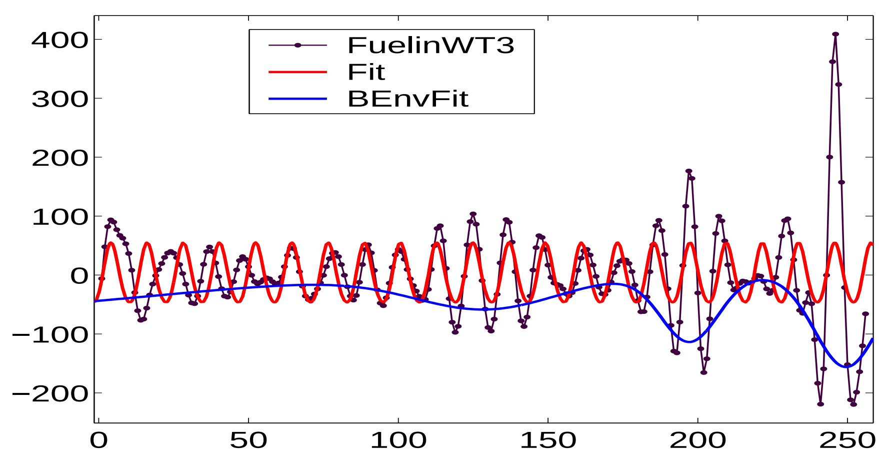


Figure 69: Third Level Bottom Envelop Fitted Using Gaussian Regression

Table 14: Third Level Lower Envelop Gaussian Regression Coefficients

n	a_n	b_n	c_n
1	-156.1	249.5	14.74
2	-112.5	197.2	13.42
3	-55.85	129.1	33.69
4	-58.82	-58.33	106.9

Figure 70 shows that the forecasting results through this process, which combines harmonic regression and Gaussian regression, are reasonable.

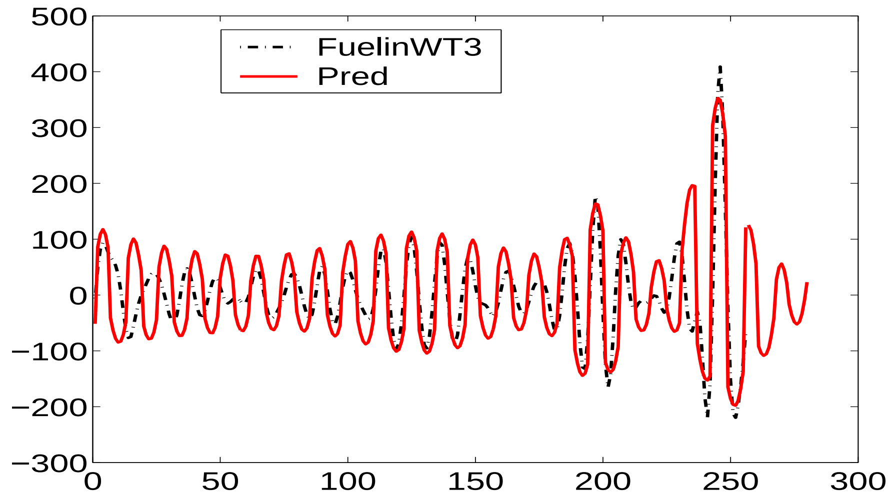


Figure 70: Third Level Forecasting Results ($\omega = 0.520$)

4.2.2.5 Fourth Level Data Analysis

The fourth level of the data in the wavelet domain represents the trend of the original data. Holt-Winters' method is used to forecast the future. The preset parameters α, β, and γ are set to be 0.1, 0.2, and 0.3, respectively. Figure 71 shows the forecasting results of Holt-Winters' method.

4.2.3 Forecasting Results

Finally, the forecasting results for the four levels in the wavelet domain are combined through the inverse WT and shown in the time domain. Figure 72 shows the forecasting results in the time domain.

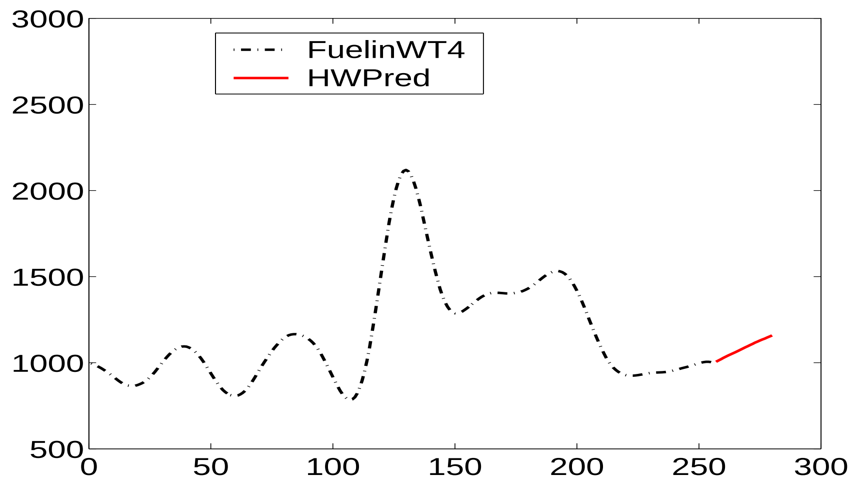


Figure 71: Fourth Level Forecasting Results by Holt-Winters' Method

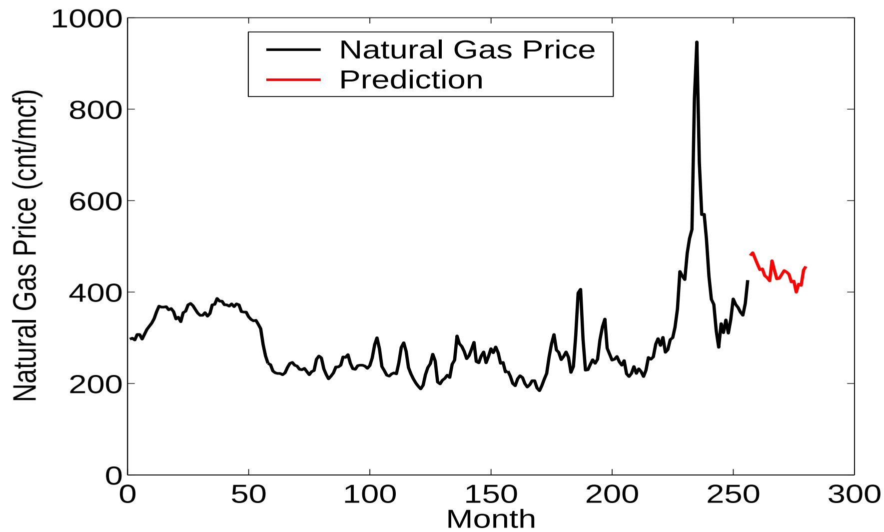


Figure 72: Forecasting Results for the Following 24 Months

4.3 Electricity Prices Forecasting

4.3.1 Historical Data

Historical data for electricity prices within a given market domain are also obtained from July 1981 through October 2002. The data set $EC = \{ec_t\}_{t=1,2,...,n}$ consists of $n = 256$ (2^8) monthly data points corresponding to $\{$(July 1981), (Aug. 1981), . . . , (Oct. 2002)$\}$. Figure 73 represents the historical data of electricity prices. The graph shows that electricity prices have a strong seasonal pattern but no apparent trend. Figure 74, which "zooms in" on electricity prices in 1986 – 1987 and 1999 – 2000, shows that electricity prices reach a maximum for each year in August and a minimum in February and that they have no obvious trend.

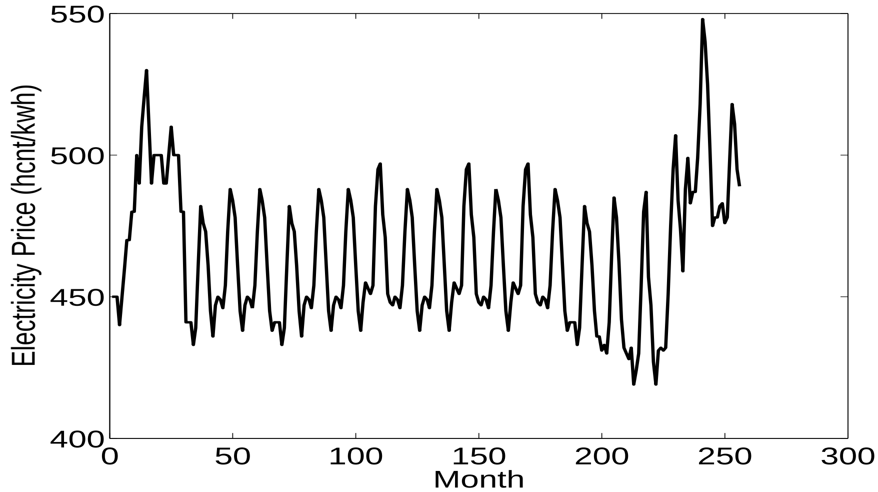


Figure 73: Electricity Industrial Sector Prices ($hcnt/kwh$)

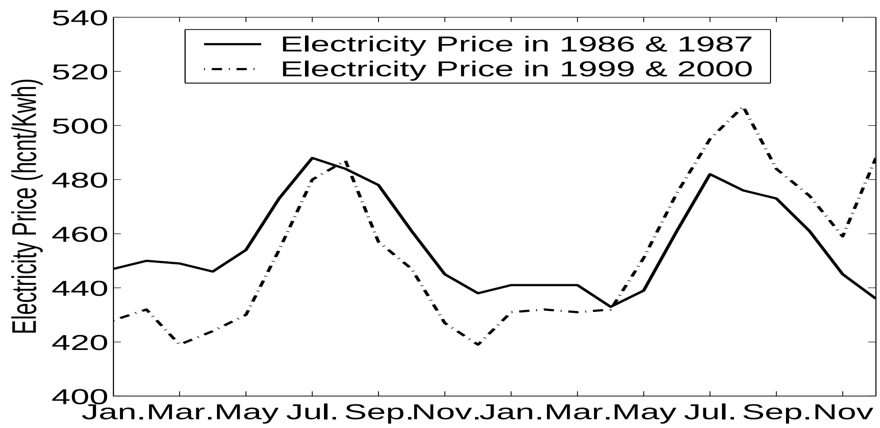


Figure 74: Seasonal Patterns Existing in the Historical Data

4.3.2 Data Analysis

4.3.2.1 Wavelet Transform

Electricity prices are transformed to the wavelet domain by the NDWT performed with the Symmlet 8 filter to extract critical information for forecasting. Figure 75 shows the electricity prices in the wavelet domain. The left graph shows the transformed data with a transform depth of five levels and the right one with a transform depth of four levels. A comparison of these two transforms shows that the four-level transform provides enough information through decomposing the data into high-frequency (first level), seasonal (second and third levels), and trend (fourth level) components. Therefore, it is used in the electricity price analysis.

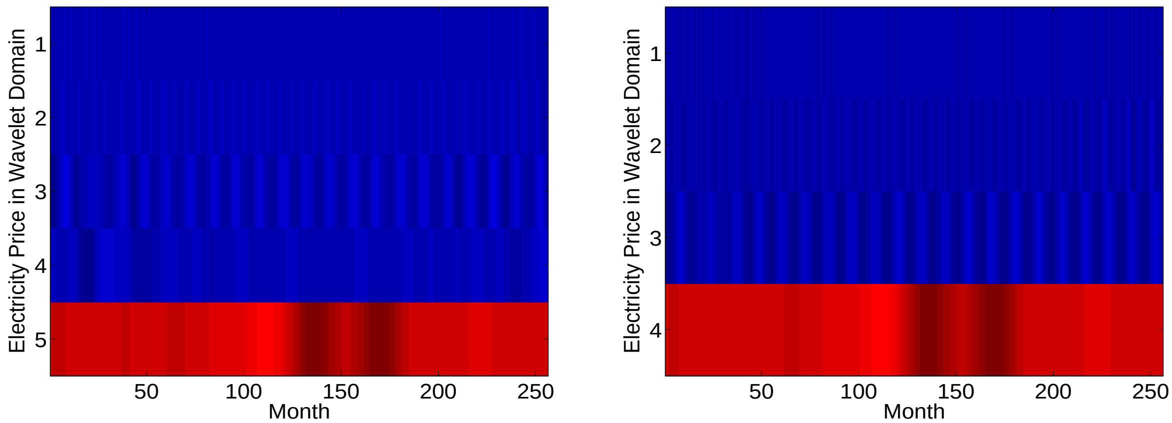


Figure 75: Electricity Prices in the Wavelet Domain, Performed with Symmlet (8)

4.3.2.2 First Level Data Analysis

The first level is shown in the left graph in Figure 76. The right part of Figure 76 plots the sample PACF together with the bounds $\pm 1.96/\sqrt{n}$. From this graph, it is easy to read off the preliminary estimator of $p = 8$.

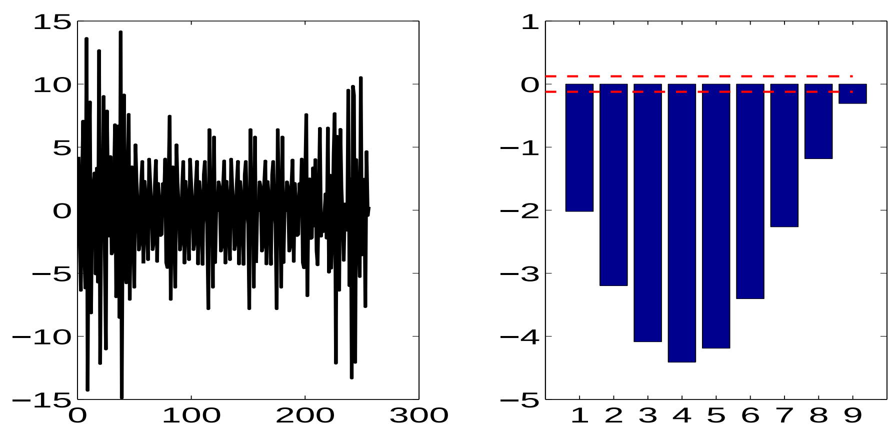


Figure 76: The First Level Data and Fitness Test

As fuel prices contribute significantly to electricity prices, a strong relationship is expected. Figure 77 plots the first level of electricity prices in the upper graph, which is treated as the output of the ARMAX model. The bottom graph shows the first level of natural gas prices, which is treated as the input to the ARMAX model. The cross correlations between these two series of data are calculated (see Figure 78). The figure shows that electricity prices are strongly correlated with fuel prices. The time lag of the external factor with respect to the output peaks at time lag zero.

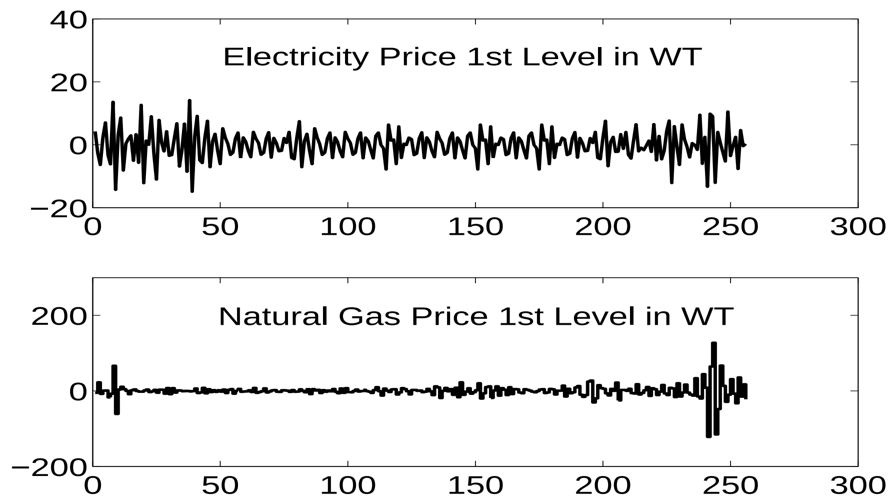


Figure 77: The First Level Data and the External Factor IDPlot

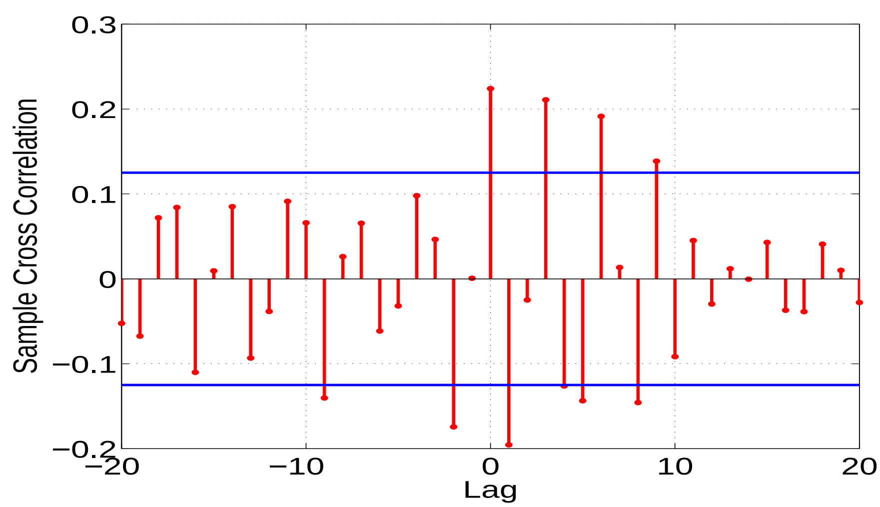


Figure 78: Correlation between the First Level Data and the External Factor

ARX[8,2,0] is used to model the first level data with

$$\hat{\phi} = \{1.857, 2.821, 3.488, 3.624, 3.307, 2.517, 1.524, 0.6724\}, \text{ and}$$

$$\hat{\xi} = \{0.02319, -0.0009138\}.$$

The upper graph of Figure 79 shows that the model residuals are not within themselves. The lower graph shows that the residuals are not correlated with the external factor. The small amplitude of the correlation functions implies that the model is good.

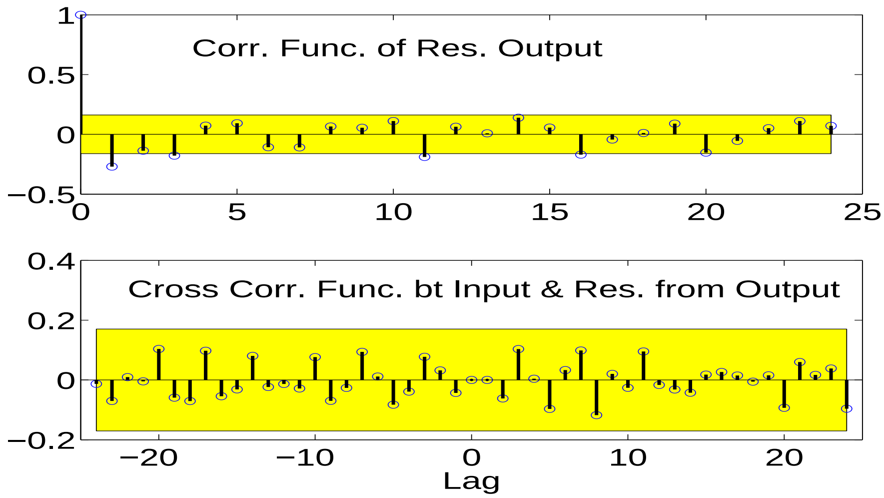


Figure 79: Correlation Relationship of the Residuals

The ARX model is then utilized to perform the forecasting. The forecasting results in Figure 80 show that the ARX model can simulate the historical data very well. Therefore, it is utilized to perform the forecasting of the following 24 months.

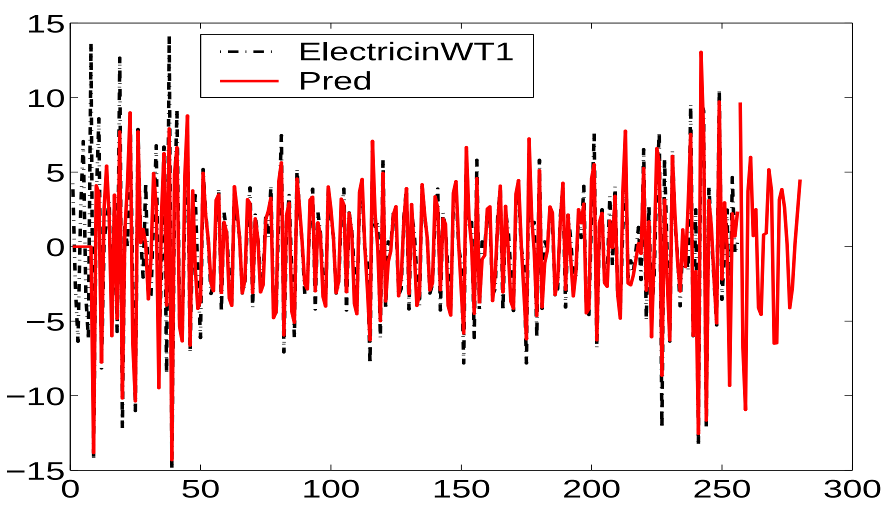


Figure 80: The First Level Data Fitted Using ARMAX Process

4.3.2.3 Second Level Data Analysis

The second level captures the seasonality in the historical data. Figure 81 shows the second level data with obvious seasonal variations with almost constant magnitude. Harmonic regression is used to account for most of the cycles present in the data. It can be seen

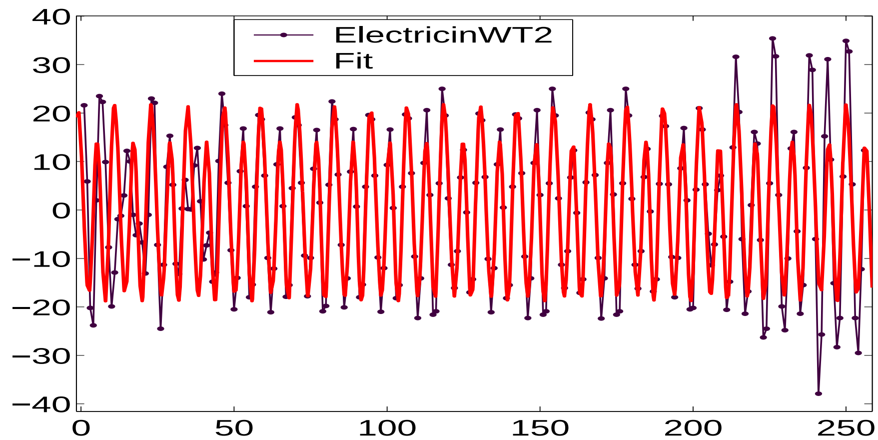


Figure 81: The Second Level Fitted Using Harmonic Regression ($\omega = 0.5254$)

that only harmonic regression is good enough to simulate the second level data. Table 15 calculates the coefficients for the harmonic regression. The frequency in the harmonic regression is 0.5254, which means that the period of the seasonal characteristics is 11.96 months. This value is very close to 12 months and the periods estimated for customer demand and natural gas prices. Figure 82 shows the forecasting results by using this method.

Table 15: Second Level Harmonic Regression Coefficients

$\omega = 0.5254$		
$\alpha_0 = 0.08684$		
n	β_n	γ_n
1	4.427	-2.845
2	11.11	-13.61
3	-2.301	1.314
4	-0.1639	0.07577
5	-0.06495	0.01435
6	-0.09834	0.009645
7	-0.07577	-0.01378
8	0.1896	0.02823

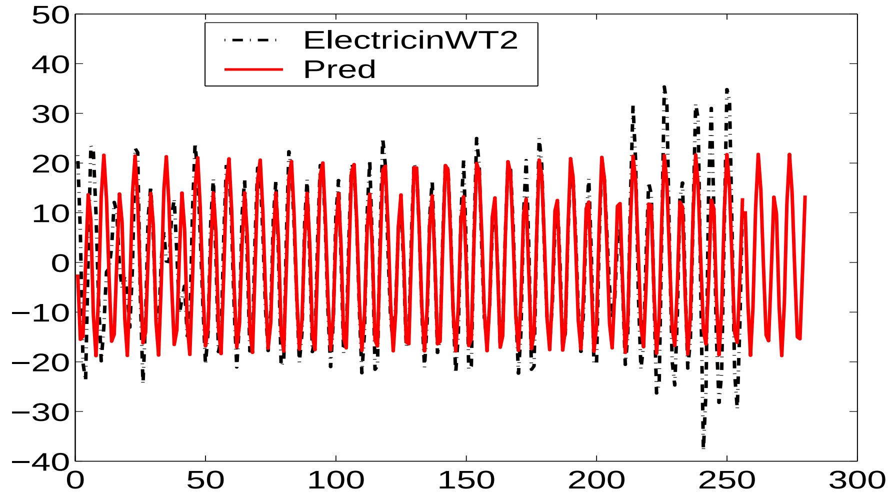


Figure 82: Second Level Forecasting Results ($\omega = 0.5254$)

4.3.2.4 Third Level Data Analysis

The third level also represents seasonal characteristics. Figure 83 shows the harmonic regression of the historical data. Table 16 gives the coefficients of these regressions. The frequency is 0.5211, which corresponds to a period of 12.06 months for the third level data. Figure 84 shows the forecasting results by utilizing harmonic regression.

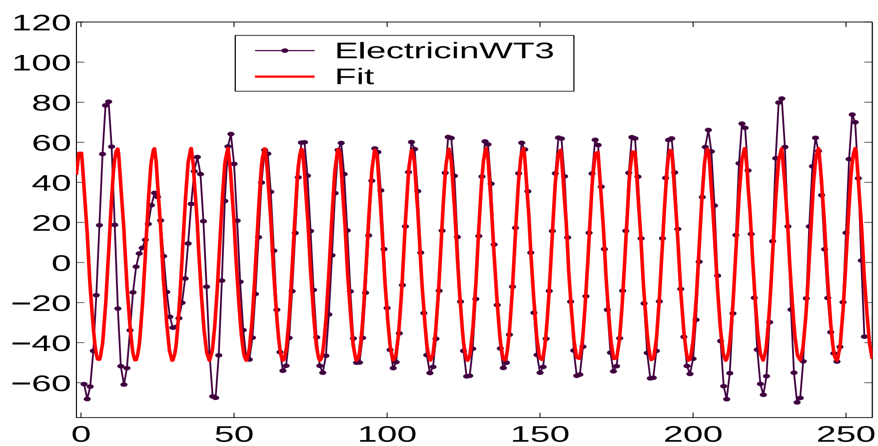


Figure 83: The Third Level Fitted Using Harmonic Regression ($\omega = 0.5211$)

4.3.2.5 Fourth Level Data Analysis

The fourth level of the data in the wavelet domain represents the trend existing in the original data. Holt-Winters' method is used to forecast the future. The preset parameters

Table 16: Third Level Harmonic Regression Coefficients

$\omega = 0.5211$		
$\alpha_0 = -0.1732$		
n	β_n	γ_n
1	51.2	-9.3
2	3.255	-2.156
3	0.641	-0.3474
4	0.2923	-0.09101
5	0.2761	-0.1199
6	0.4395	-0.03822
7	0.4219	0.2321
8	0.2239	0.3051

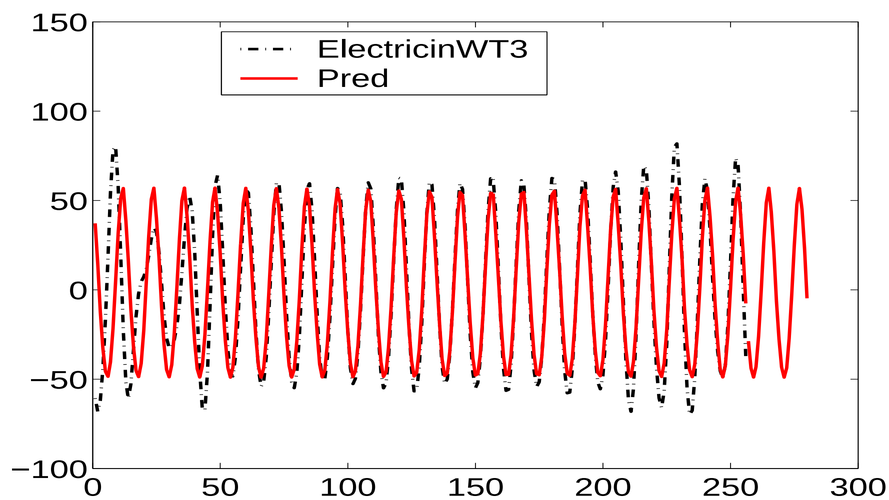


Figure 84: Third Level Forecasting Results ($\omega = 0.5211$)

α, β, and γ are set to be 0.1, 0.2, and 0.3, respectively. Figure 85 shows the forecasting results of Holt-Winters' method.

4.3.3 Forecasting Results

Finally, the forecasting results for the four levels in the wavelet domain are combined through the inverse WT and shown in the time domain. Figure 86 shows the forecasting results in the time domain. The following 24 months of forecasting data exhibit the seasonal characteristics identified in the historical data. The magnitude shows a slow increasing trend.

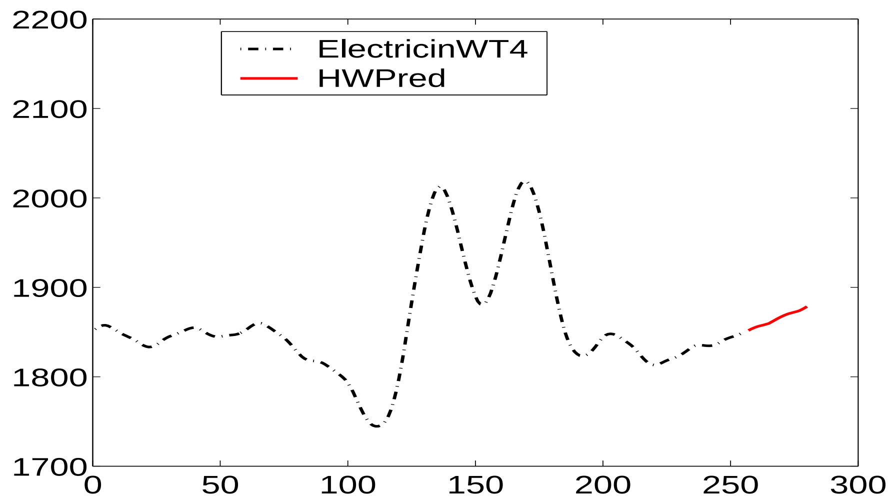


Figure 85: Fourth Level Forecasting Results by Holt-Winters' Method

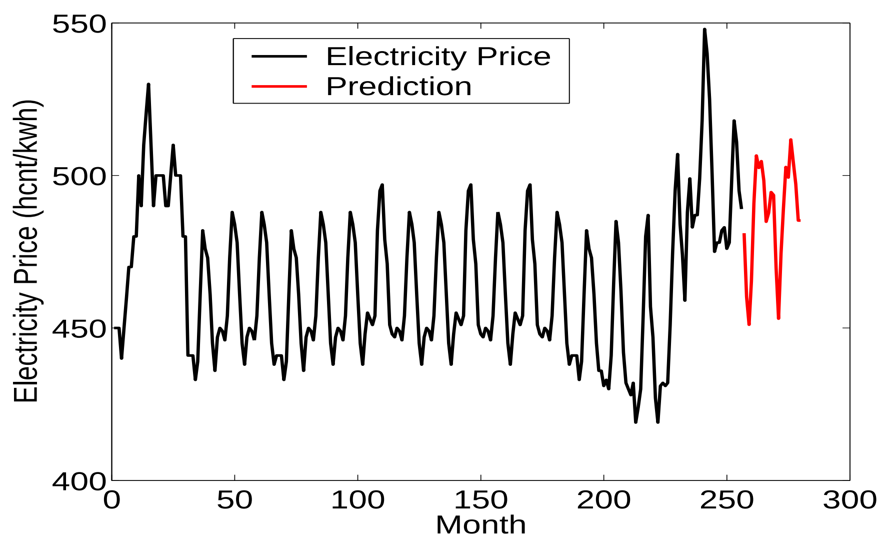


Figure 86: Forecasting Results for the Following 24 Months

4.4 Forecasting Errors

Forecasting errors are calculated through a comparison with the real data obtained from the electric market. Figures 87, 88, and 89 show differences between the forecasting results and the real values. The forecasting results for electric prices and customer demand behave much better than those for natural gas prices. The high volatility in the recent business environment and government input contribute to the gap between the forecasts and the real data. Different measurements of the forecasting errors, mean squared error (MSE), mean absolute deviation (MAD), mean absolute percentage error (MAPE), and bias, for natural gas prices, electricity prices, and customer demand are shown in Table 17. The MSE can be related to the variance of the forecast errors. The MAD can be used to estimate the standard deviation of the forecast errors assuming that the forecast errors are normally distributed. MAPE is the average absolute error as percentage of the real value of the forecasting variable. Bias determines whether a forecast method consistently over- or underestimates the forecasting variable.

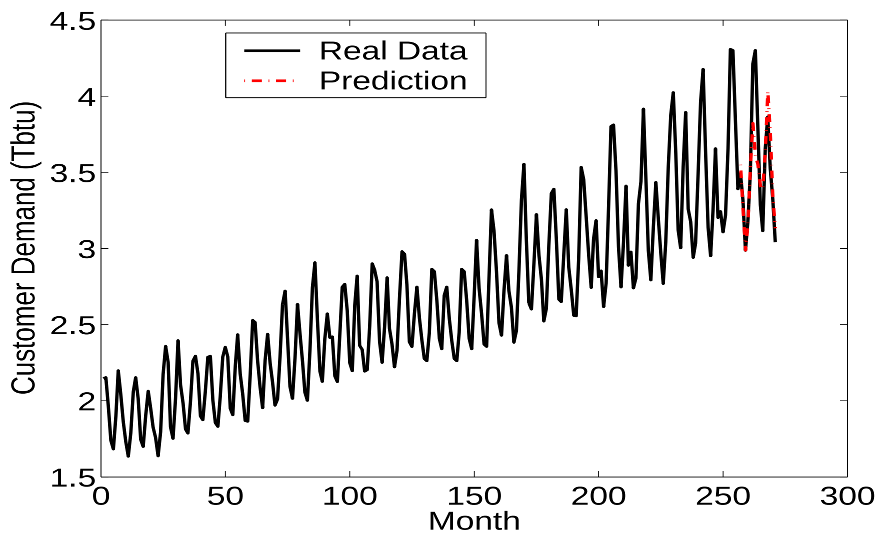


Figure 87: Customer Demand Validation ($Tbtu$)

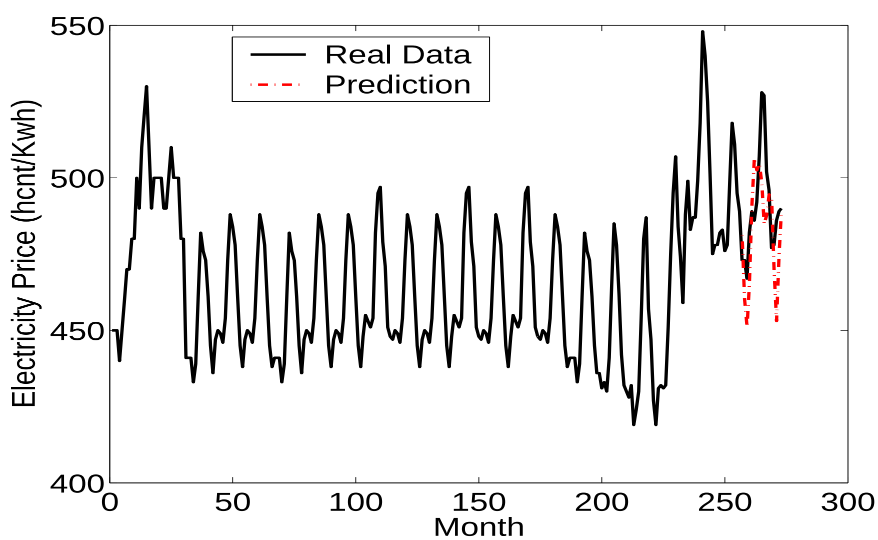


Figure 88: Electricity Price Validation ($hcnt/kwh$)

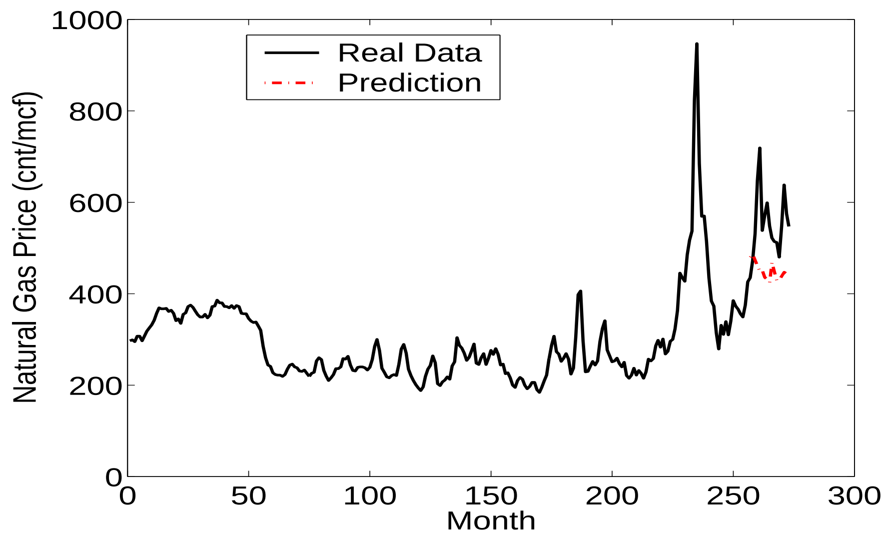


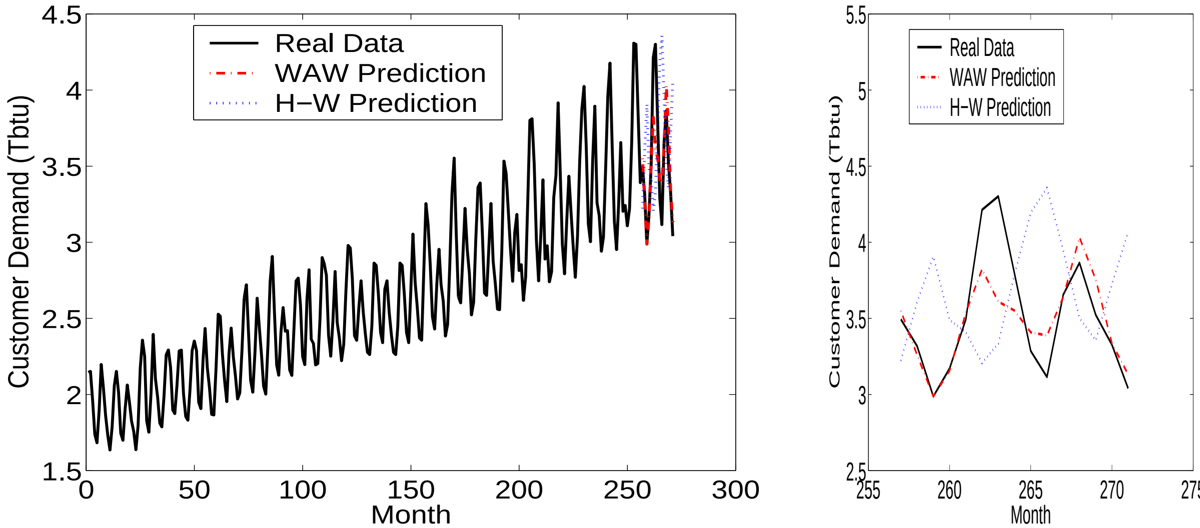
Figure 89: Natural Gas Price Validation (cnt/mcf)

Table 17: Forecasting Errors

	MSE	MAD	MAPE	Bias
Natural Gas Prices	16273	110.5897	19.0209	1762.6
Electricity Prices	341.5384	14.6856	2.9655	134.6995
Customer Demand	0.0585	0.1606	4.2645	0.4193

4.5 Comparisons With Holt-Winters' Method

Holt-Winters' method is usually used in engineering for performing forecasting for historical data with level, trend, and seasonality. It is applied to the historical data of customer demand, natural gas prices, and electricity prices. The forecasting results from Holt-Winter's method are compared with the forecasting results from the WAW method and the actual data. Figures 90, 91, and 92 show the comparisons between these three sets of data for customer demand, natural gas prices, and electricity prices, respectively. The results demonstrate that the WAW method can better simulate the impact of the external business environment on the evolution of forecasting, and thus lead to more accurate overall forecasting. The right graph of Figure 90 "zooms in" on customer demand for the forecasting period. From this figure, it is shown that the WAW method can better account for the seasonal characteristics in the historical data and provide more accurate forecasts.



(Customer Demand) ("Zoom In")

Figure 90: Residential and Commercial Demand ($Tbtu$)

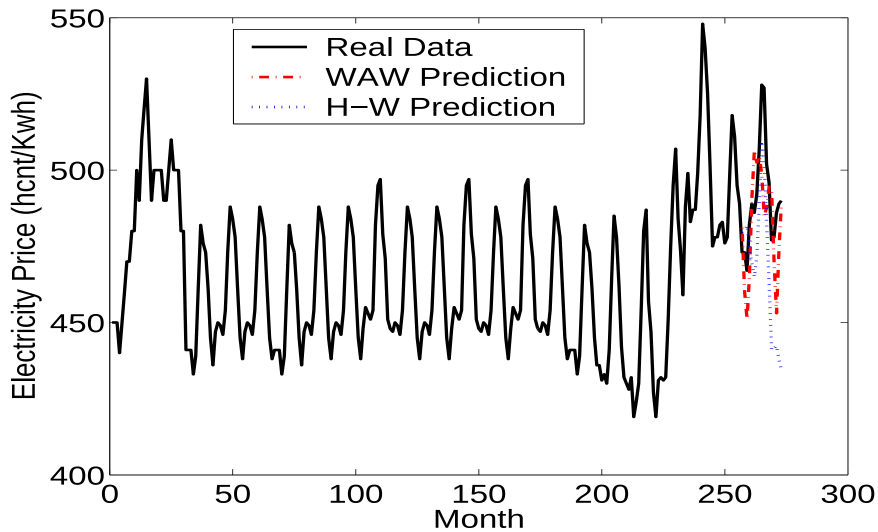


Figure 91: Electricity Price Comparison ($hcnt/kwh$)

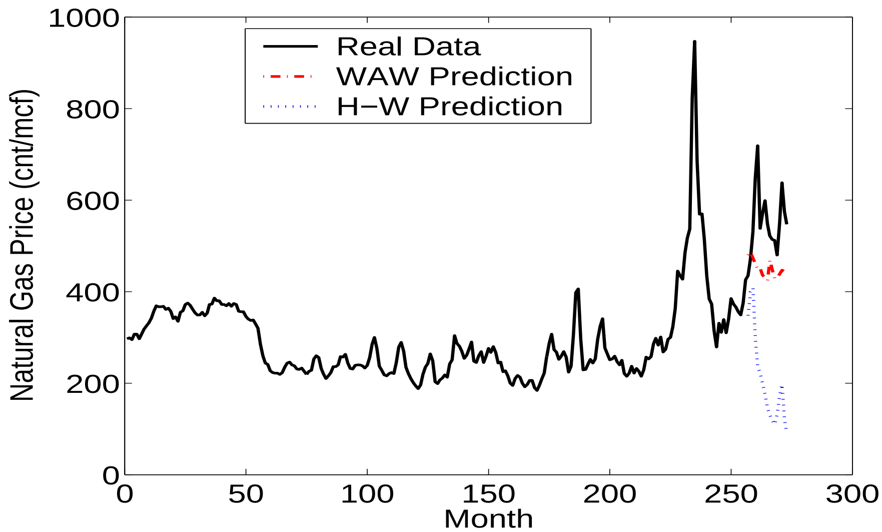


Figure 92: Natural Gas Price Comparison (cnt/mcf)

CHAPTER V

POWER PLANT FLEET MANAGEMENT

The DM process considering cross-scale interactions was carried out to identify the optimal SOS, SMS, and further SCEP to achieve system excellence under normal operating conditions. Then scenario analysis was utilized to describe the evolutions of the power plant under different environments.

5.1 Unit Conditions and System Characteristics

5.1.1 Unit Load Settings

The typical power plant to which the DM process is applied owns five generation units. The operation is discretized into five conditions based on the power output. The production at each operating condition for each generation unit is given in Table 18. Data are normalized by the highest unit production ($\mathcal{HUP}$). In Table 18, the first and fifth generation units produce the highest output at their peak load operating conditions. The outputs at other operating conditions for these two units and the outputs for all the other generation units are normalized by the value of $\mathcal{HUP}$.

Table 18: Normalized Generation Unit Output

Unit	Part Load ($\mathcal{HUP}$)	Base Load ($\mathcal{HUP}$)	Peak Load ($\mathcal{HUP}$)	Maintenance	Off
1	0.6137	0.7659	1.00	0	0
2	0.5962	0.7484	0.9912	0	0
3	0.5787	0.7309	0.9825	0	0
4	0.5962	0.7484	0.9912	0	0
5	0.6137	0.7659	1.00	0	0

5.1.2 System Capacity

System capacity is determined based on the number of generation units that are committable and the conditions they are operating at. In this study, the system capacity is defined to be the total system output if all the committable generation units are operating at their base loads. SAC is determined to be 80% of the system capacity. Table 19 shows the system capacity and SAC.

Table 19: System Capacity and Available Capacity

Unit	Part Load	**Base Load**	Peak Load	Maintenance	Off
1	0.6137	**0.7659**	1.00	0	0
2	0.5962	**0.7484**	0.9912	0	0
3	0.5787	**0.7309**	0.9825	0	0
4	0.5962	**0.7484**	0.9912	0	0
5	0.6137	**0.7659**	1.00	0	0
System Capacity = **3.7595** $\mathcal{HUP}$					
SAC = **3.008** $\mathcal{HUP}$					

5.1.3 Economical Operating Period

The EOP of a system is the period of time that the power plant can focus on minimizing LCCs. For long-term planning, this value can not be determined once and then utilized in all cases. As a remote target approaches, forecasting information becomes more accurate, and then the EOP should be updated. The first estimate of the EOP is shown in Figure 93.

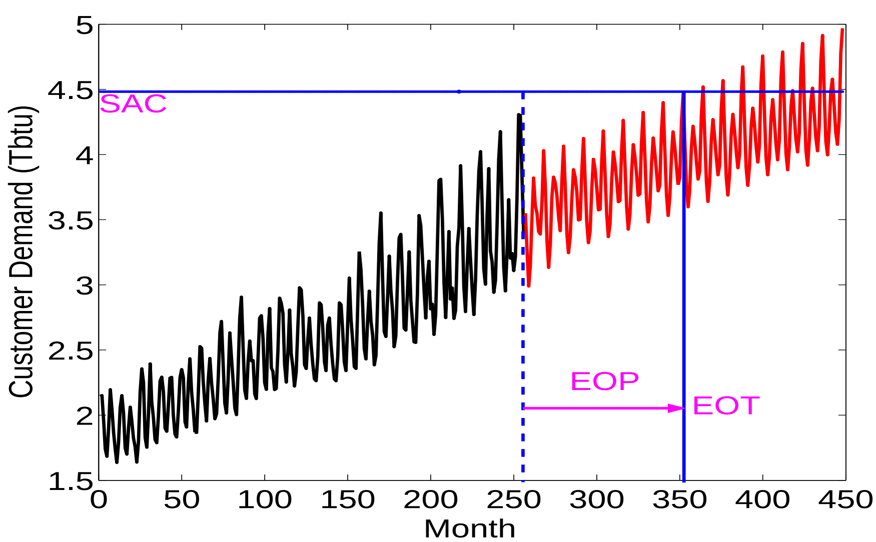


Figure 93: Economical Operating Period

Figure 93 shows that when the SAC meets customer demand, the EOP is determined.

$$\text{EOP} = 32 \text{ Quarters}$$

The EOT is the point in time at which the EOP ends. It depends on the starting point of the planning process. If the starting point of the power plant planning is Nov. 2002, EOT can thus be determined as

$$\text{EOT} = \text{Nov. 2010 Yr.}$$

5.1.4 Operation Profile

Table 20 illustrates the operating profile that each generation unit adopts. Based on the continuous operating profile, the maintenance factors are determined by normalizing the FFH with the actual operating hours for combustor, hot-gas path, and major inspections. FFH consider the specifics of the continuous duty cycles relating to fuel type, load setting, and steam or water injection. The determination of the FFH for combustor, hot-gas path, and major inspections are determined based on the operating profile provided in the table.

Table 20: Continuous Operation Profile

Operation	Coninuous
Hot Start (Down <4 Hr.)	10%
Warm 1 Start (Down 4 – 20 Hr.)	5 %
Warm 2 Start (Down 20 – 40 Hr.)	5%
Cold Start (Down > 40 Hr.)	80%
Hours/Start	400
Hours/Year	8200
Starts/Year	21
Percent Trips	20%
Number of Trips/Year	4

5.1.5 Operating Condition Ranking

The operating conditions for each generation unit are not equally efficient. From an economic aspect, the selection of operating conditions for each generation unit to meet the forecasted customer demand will significantly affect LCCs, especially fuel costs. Select the most efficient available operating conditions for each generation unit in order to minimize the total cost is one significant step to operate the whole power plant. This necessitates the need to rank the operating conditions for each generation unit and for the whole power plant. The ranking criterion defined as the ratio of output of a unit at a certain operating condition for a given period of time to the FFH for that period of time is calculated and listed in Table 21.

Table 21: Operating Condition Ranking

Unit	Part Load	Base Load	Peak Load	Maintenance	Off
1	14	12	4	0	0
2	8	10	2	0	0
3	6	7	1	0	0
4	9	11	3	0	0
5	15	13	5	0	0

Table 21 shows the rankings for all the operating conditions for all generation units. For each generation unit, the most efficient available operating condition will be first selected to satisfy customer demand. The increases in customer demand will require increasing the load levels of the generation unit that has the most efficient operating condition to provide the production.

5.2 System Operating Strategies And System Maintenance Schedules

Different colors are used to represent different operating conditions for each generation unit in order to make the system status easily presentable. Table 22 shows the relationship between the operating condition and its color. The combination of these five colors represents the system status. The following combination in Figure 94 shows that the first generation unit is operating at part load, the second is operating at base load, the third is under maintenance, the fourth is at peak load, and the fifth is in an off condition.

Figure 94: System Status vs. Color

Table 22: Operating Condition vs. Color

Part Load	Base Load	Peak Load	Maintenance	Off
Yellow	Green	Magenta	Red	Blue

5.2.1 Baseline SMS and SOS

The baseline condition is defined as the condition in which the power plant operates according to the determined SOS for each quarter. The recommended SMS can be carried out perfectly. No unscheduled events, such as unscheduled maintenance, unexpected customer demand, and so forth, interrupt the operating process. The forecasting information provided is based on normal economic development, normal weather conditions, and no special events.

5.2.1.1 SMS and SOS

The baseline operation of the whole power plant can be illustrated in Figure 95. For each quarter, the system status will be updated. Each generation unit operates according to the new system status so that the system output can satisfy customer demand at a minimal cost. The far left column in Figure 95 shows the system status selected for each quarter. It appears that for several quarters the system status remains the same due to the fact that customer demand varies very slowly, so the previous system status still remains optimal.

The right part of Figure 95 shows the SMS. The horizonal direction displays the weekly activities of the power plant. If a quarter is free from scheduled maintenance activities for each week, each generation unit follows the operating condition determined at the beginning of that quarter. The corresponding row will be blank, e.g., during entire 6^{th} quarter, no maintenance activities take place. The power plant operates the same for all the weeks in the quarter. The occurrence of the "point events" triggers the switches of the operating conditions for each generation unit. At the spot in the figure of the corresponding week, the operating conditions for all the generation units are given. After the "point event" is resolved, the system status will recover to the one that was selected for the current quarter. For example, in the 4^{th} quarter, the system status is base load for all the generation units, but in the 6^{th} week, the system status has been changed to base load for units 1, 2, 3, and 5 because unit 4 is under scheduled maintenance. In the 7^{th} week, the system status needs to be changed to base load for units 1, 3, 4 and 5 because unit 2 requires scheduled maintenance.

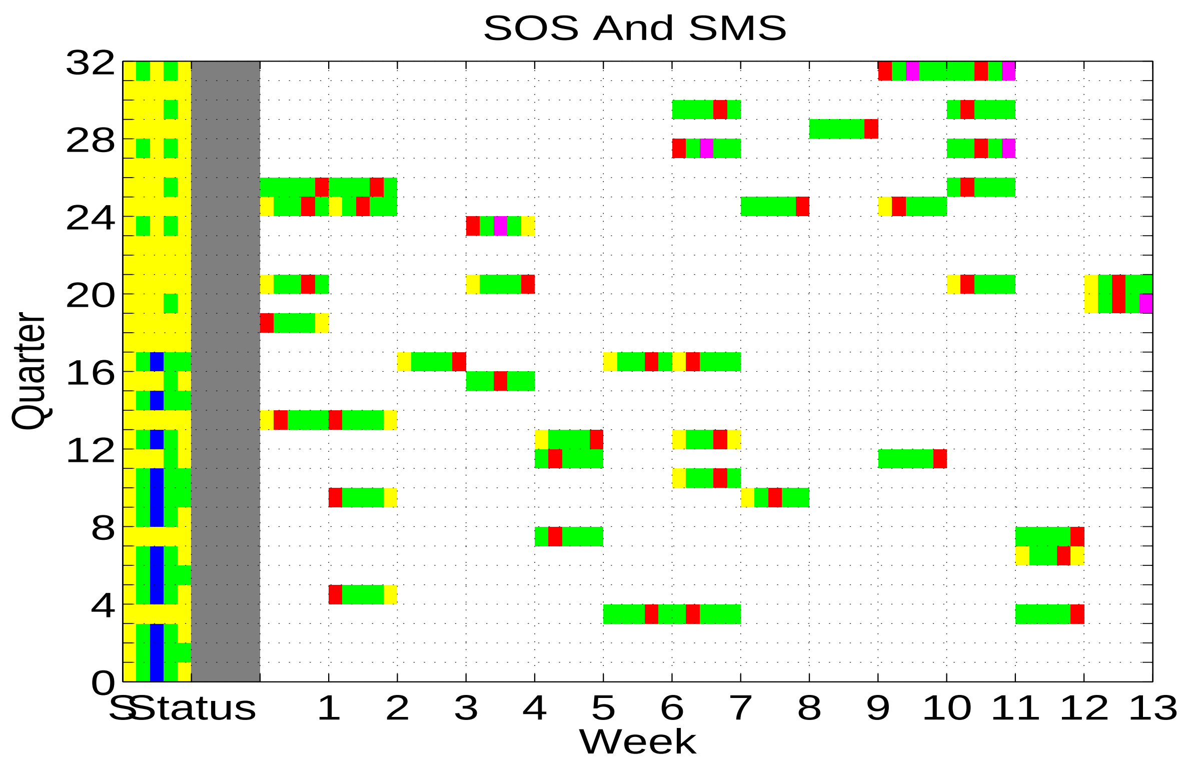


Figure 95: Baseline: SOS and SMS

5.2.1.2 System Production vs. Customer Demand

The objective of updating the system status per quarter is to provide customer demand at a minimal cost, without extra expenditures on too much power generation or penalties for curtailing customer demand. Figure 96 shows customer demand and the system generation based on the system status selected for each quarter, shown in Figure 95. From this figure, it can be seen that the power plant can well satisfy the forecasted customer demand while capturing variations in customer demand.

In cases in which scheduled maintenance occurs, system generation will decrease because the generation unit that was taken out of service for maintenance. Table 23 illustrates the

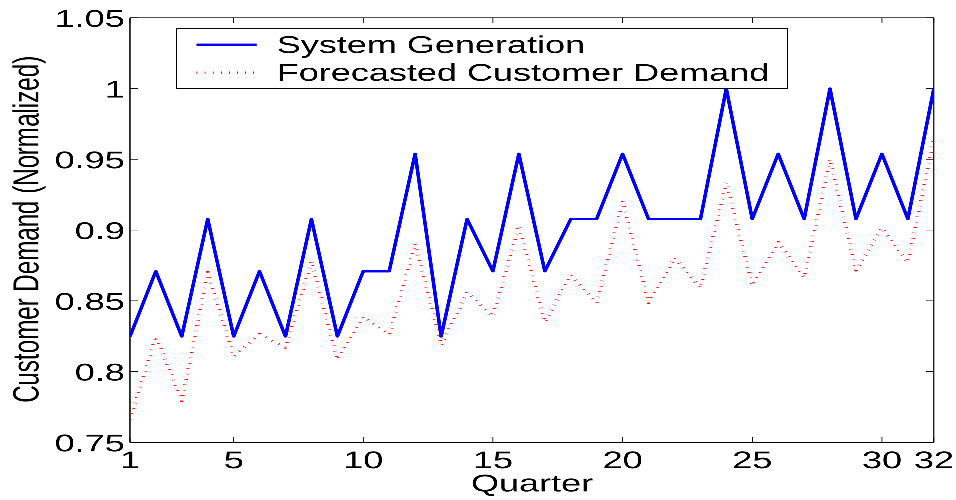


Figure 96: Baseline: System Generation vs. Customer Demand

maintenance activities in the 4^{th} quarter. Unit 4 requires maintenance in the 6^{th} week, unit 2 requires maintenance in the 7^{th} week, and unit 5 requires maintenance in the 12^{th} week. All these maintenance activities are scheduled based on the cumulative FFH. Figure 97 compares the system generation under two conditions: one is if the system remains in the same operating status and the other is to switch to a new one during the maintenance window, and compares those with customer demand. It can be seen that if system status is not adjusted, due to the scheduled maintenance, the system cannot meet customer demand in the maintenance window. By switching the load levels of other generation units, the system is able to meet customer demand. For example, in the 6^{th} week, all the other generation units except the one in maintenance have switched their operating conditions from part load to base load to compensate the loss of generation due to the scheduled maintenance. The changes in the load levels are shown in Table 24.

Table 23: Baseline: Maintenance Activities in the 4^{th} Quarter

Week	Unit	Maintenance Type
6	4	S
7	2	S
12	5	S

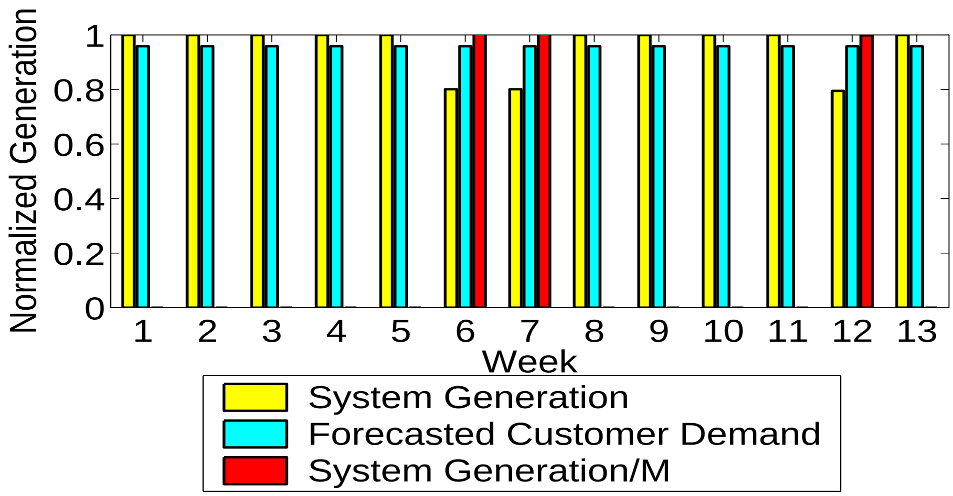


Figure 97: Baseline: System Reactions in the 4^{th} Quarter

Table 24: Baseline: System Status Adjustments in the 4^{th} Quarter

Unit	Before	6^{th} Week	7^{th} Week	12^{th}Week
1	PartLoad	BaseLoad	BaseLoad	BaseLoad
2	PartLoad	BaseLoad	Maintenance	BaseLoad
3	PartLoad	BaseLoad	BaseLoad	BaseLoad
4	PartLoad	Maintenance	BaseLoad	BaseLoad
5	PartLoad	BaseLoad	BaseLoad	Maintenance

Another example is the operation during the 14^{th} quarter. Table 25 gives the maintenance activities in this quarter. Figure 98 compares the system generation under two conditions: one is if the system remains in the same operating status and the other is to switch to a new one during the maintenance window, and compares those with customer demand. Table 26 shows the load level changes when scheduled maintenance activities occur.

Table 25: Baseline: Maintenance Activities in the 14^{th} Quarter

Week	Unit	Maintenance Type
1	2	S
2	1	S

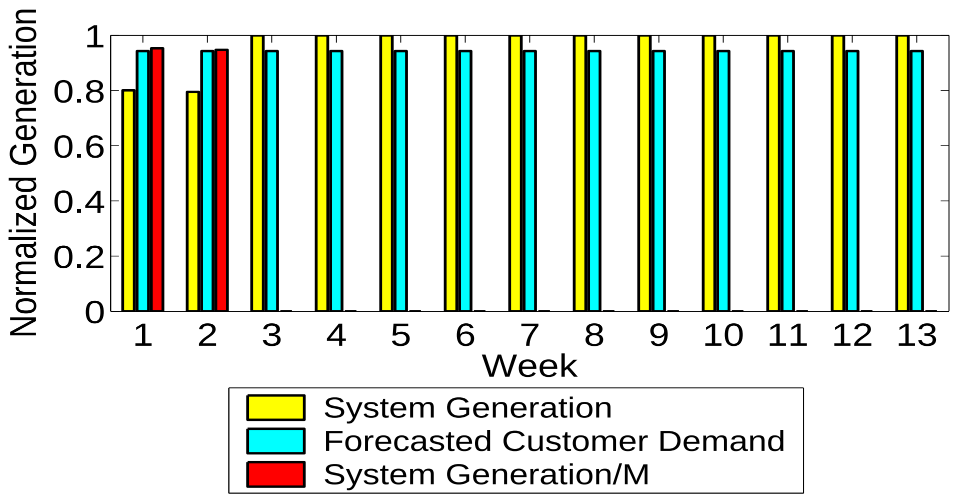


Figure 98: Baseline: System Reactions in the 14^{th} Quarter

Table 26: Baseline: System Status Adjustments in the 14^{th} Quarter

Unit	Before	1^{st}Week	2^{nd}Week
1	PartLoad	PartLoad	Maintenance
2	PartLoad	Maintenance	BaseLoad
3	PartLoad	BaseLoad	BaseLoad
4	PartLoad	BaseLoad	BaseLoad
5	PartLoad	BaseLoad	PartLoad

5.2.1.3 Life Cycle Cost

Fuel and maintenance costs are two major cost components of the total LCC of the power plant operation. The maintenance cost is closely related to the maintenance activities, including startup costs, shutdown costs, material costs, downtime costs, labor fees, and electricity purchase costs, if necessary. The fuel cost is determined mainly by the system generation. Figure 99 shows the fuel cost, maintenance cost, and total cost distributions over the EOP. Clearyly, maintenance activities have contributed to the higher cost, such as the 4^{th} and 21^{th} quarters. The total LCC is $3.6821\mathcal{NV}$, where $\mathcal{NV}$ is the value used to normalize the total cost for the baseline operation.

In reality, this baseline SOS and SMS can seldom be carried out due to various factors that act as a trigger that diverts the system status from the ideal one. Thus, how the power plant reacts under various situations is of interest in this study. The next section

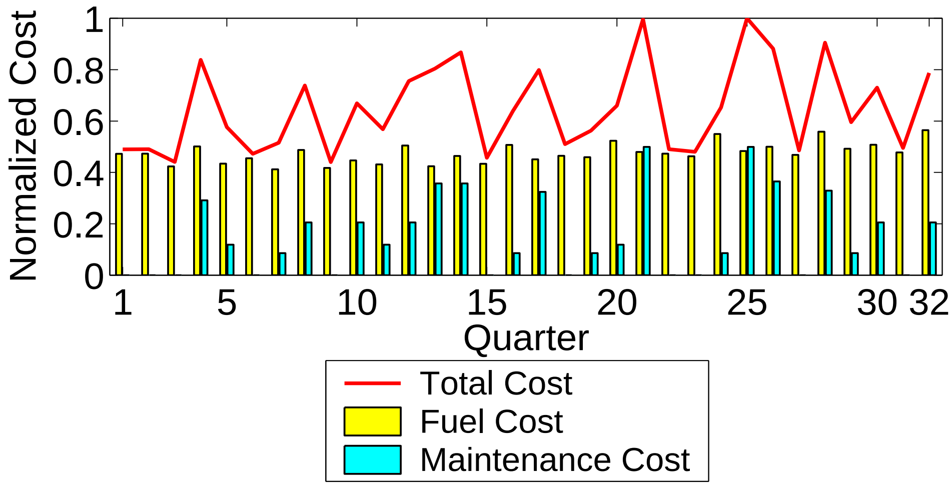


Figure 99: Baseline: Power Plant Cost Distributions

will discuss these situations and the responses of the system to them. Figure 100 shows the locations at which these deviations may occur. The changes in SOS, SMS, and cost distributions associate with fuel costs, maintenance costs, and total costs will be discussed in detail.

5.2.2 Deviation Analysis

5.2.2.1 Deviation 1

In this case, an unscheduled maintenance occurs when no scheduled maintenance has been planned. Table 27 shows that the unscheduled maintenance for unit 1 occurs in the 9^{th} week of the 4^{th} quarter and lasts for 2 weeks. Since no scheduled maintenance was planned in the baseline operation, there is no conflict in maintenance resource allocation. Figure 101 shows how the power plant will operate when this unscheduled maintenance occurs.

Table 27: Deviation 1: Unscheduled Maintenance

Quarter	Week	Unit	Duration (Week)
4	9	1	2
NO CONFLICTS			

Table 28 illustrates the maintenance activities, including scheduled and unscheduled maintenance in the 4^{th} quarter. Figure 102 compares the system generation under two

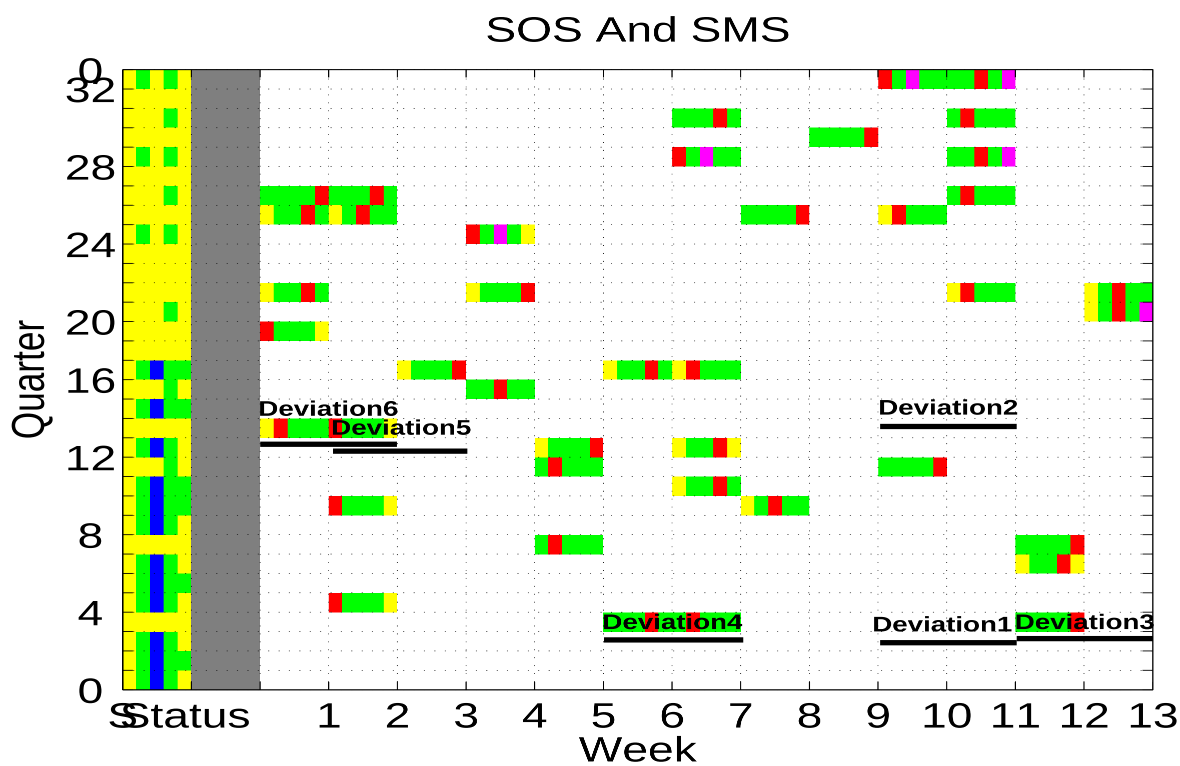


Figure 100: Deviation Locations in the Baseline Operation

conditions: one is if the system remains in the same operating status and the other is to switch to a new one during the maintenance window, and compares those with customer demand. If system status is not adjusted during the maintenance window, customer demand cannot be satisfied because some generation units have been taken offline for maintenance. By switching the load level of other available generation units according to Table 29, the system is able to generate enough power to meet customer demand.

Table 28: Deviation 1: Maintenance Activities in the 4^{th} Quarter

Week	Unit	Maintenance Type
6	4	S
7	2	S
12	5	S
9	1	U
10	1	U

The introduction of unscheduled maintenance has an impact on the later maintenance

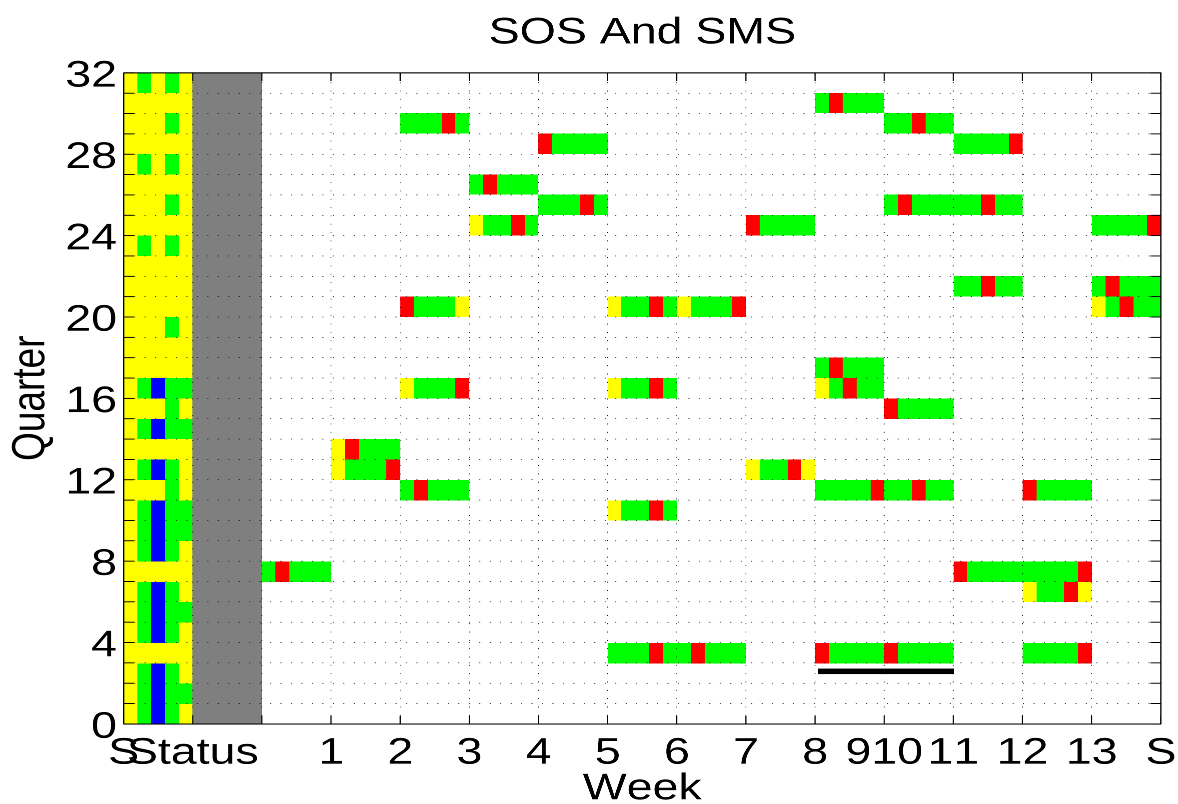


Figure 101: Deviation 1: SOS and SMS

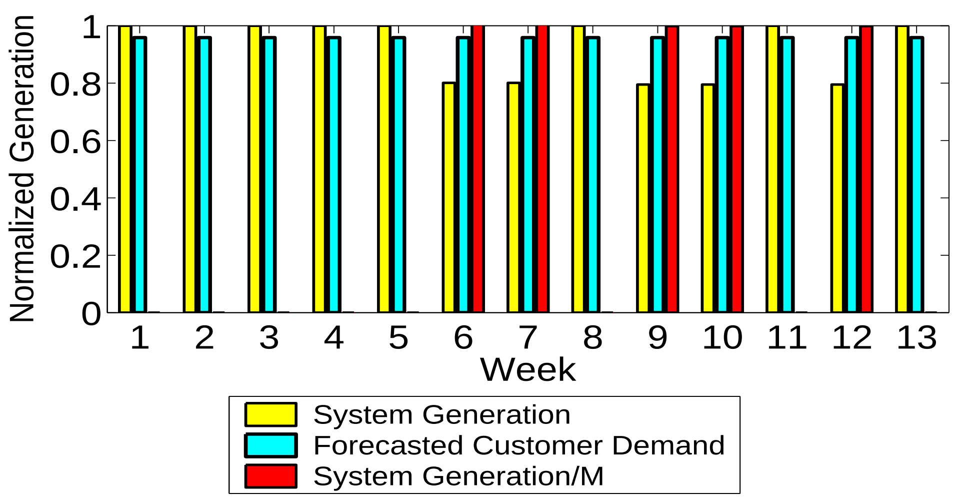


Figure 102: Deviation 1: System Reactions in the 4^{th} Quarter

activities of the power plant. Table 30 gives the maintenance activities in the 14^{th} quarter. Only one scheduled maintenance has been planned for unit 2 in the 2^{nd} week. The scheduled maintenance for unit 1 in the 1^{st} week recommended in the baseline operation is not needed

Table 29: Deviation 1: System Status Adjustments in the 4^{th} Quarter

Unit	Before	6^{th} Week	7^{th} Week	12^{th} Week	9^{th} Week	10^{th} Week
1	PartLoad	BaseLoad	BaseLoad	BaseLoad	Maintenance	Maintenance
2	PartLoad	BaseLoad	Maintenance	BaseLoad	BaseLoad	BaseLoad
3	PartLoad	BaseLoad	BaseLoad	BaseLoad	BaseLoad	BaseLoad
4	PartLoad	Maintenance	BaseLoad	BaseLoad	BaseLoad	BaseLoad
5	PartLoad	BaseLoad	BaseLoad	Maintenance	BaseLoad	BaseLoad

because of the earlier unscheduled maintenance. Figure 103 compares the system generation under two conditions and then compares them with customer demand. Table 31 shows how the load level changes when maintenance activities occur.

Table 30: Deviation 1: Maintenance Activities in the 14^{th} Quarter

Week	Unit	Maintenance Type
2	2	S

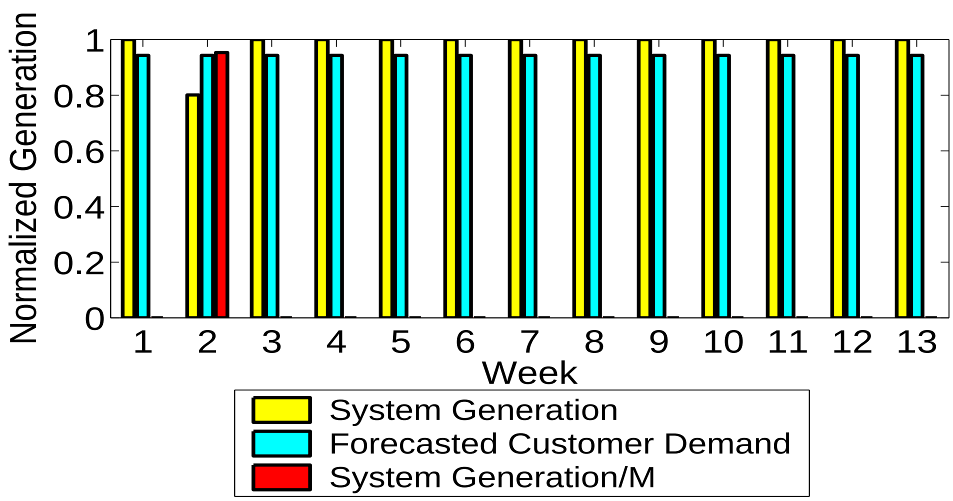


Figure 103: Deviation 1: System Reactions in the 14^{th} Quarter

The distributions of fuel costs, maintenance costs, and total costs over the EOP are shown in Figure 104. Total costs peak in the 4^{th} quarter because maintenance activities cause maintenance costs to rise sharply. However, after the 4^{th} quarter, the total cost distribution becomes more smooth.

Table 31: Deviation 1: System Status Adjustments in the 14^{th} Quarter

Unit	Before	2^{nd}Week
1	PartLoad	PartLoad
2	PartLoad	Maintenance
3	PartLoad	BaseLoad
4	PartLoad	BaseLoad
5	PartLoad	BaseLoad

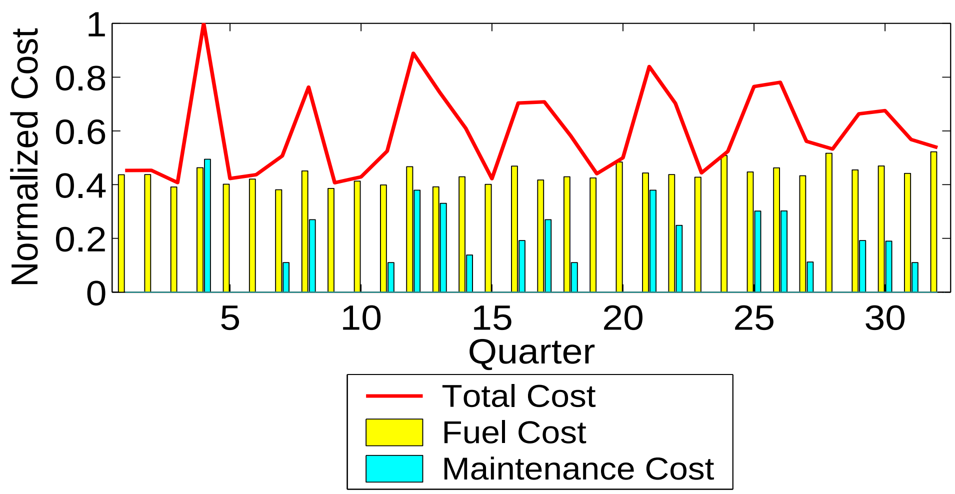


Figure 104: Deviation 1: Power Plant Cost Distributions

5.2.2.2 Deviation 2

An unscheduled maintenance takes place when no scheduled maintenance has been planned at a specific time. However, in this case, the unscheduled maintenance occurs much later in the EOP than it did in Deviation 1, in the 14^{th} quarter. Table 32 shows the condition under which the unscheduled maintenance occurs and Figure 105 shows its effect on the operating of the power plant.

Table 32: Deviation 2: Unscheduled Maintenance

Quarter	Week	Unit	Duration (Week)
14	9	1	2
NO CONFLICTS			

Table 33 gives the maintenance activities in this quarter. Figure 106 compares the

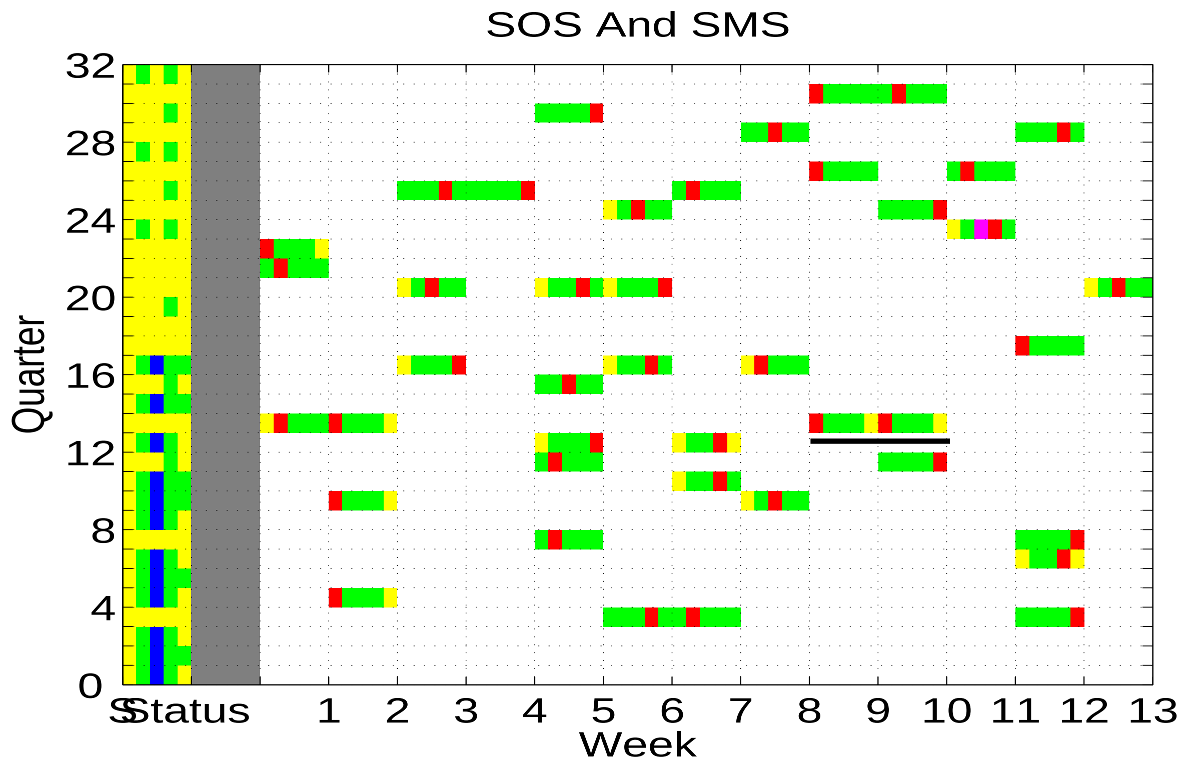


Figure 105: Deviation 2: SOS and SMS

system generation if the maintenance activity is ignored with the system generation if the system status is adjusted during the maintenance window, and compares them with customer demand. Table 34 shows the load level changes when maintenance activities occur.

Table 33: Deviation 2: Maintenance Activities in the 14^{th} Quarter

Week	Unit	Maintenance Type
1	2	S
2	1	S
9	1	U
10	1	U

Figure 107 shows the distributions of fuel costs, maintenance costs, and total costs over the EOP. This time total costs peak in the 14^{th} quarter due to the introduction of an unscheduled maintenance. The total cost of this case is 1.0142 times the total cost incurred in Deviation 1. This means that the later an unscheduled maintenance occurs, the harder

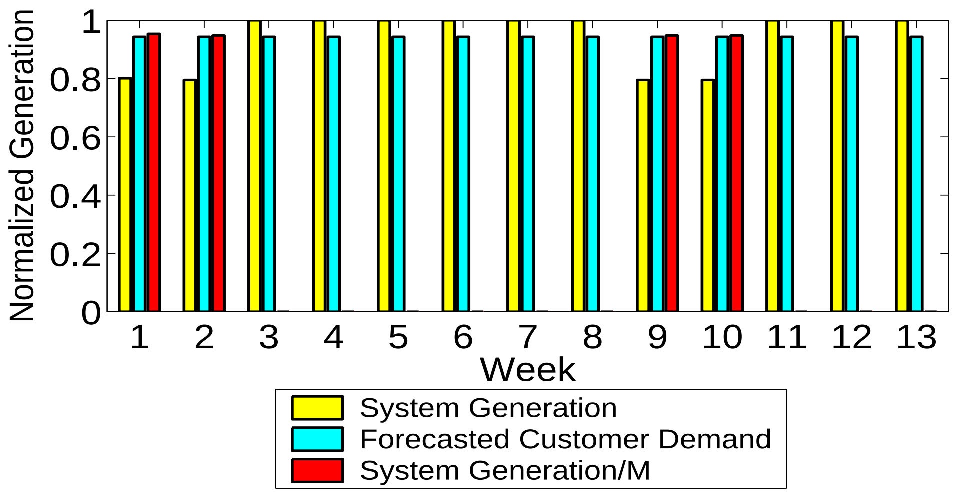


Figure 106: Deviation 2: System Reactions in the 14^{th} Quarter

Table 34: Deviation 2: System Status Adjustments in 14^{th} Quarter

Unit	Before	1^{st}Week	2^{nd}Week	9^{th}Week	10^{th}Week
1	PartLoad	PartLoad	Maintenance	Maintenance	Maintenance
2	PartLoad	Maintenance	BaseLoad	BaseLoad	BaseLoad
3	PartLoad	BaseLoad	BaseLoad	BaseLoad	BaseLoad
4	PartLoad	BaseLoad	BaseLoad	BaseLoad	BaseLoad
5	PartLoad	BaseLoad	PartLoad	PartLoad	PartLoad

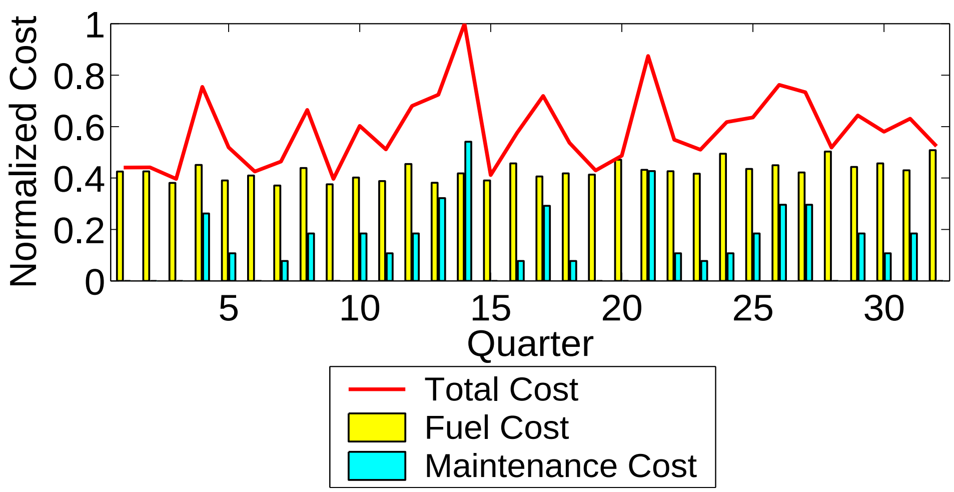


Figure 107: Deviation 2: Power Plant Cost Distributions

and more expensive it is to remedy the loss of generation due to higher customer demand needed to satisfy.

5.2.2.3 Deviation 3

In this case, an unscheduled maintenance takes place when a scheduled maintenance in the baseline operation has been planned. Table 35 shows the condition under which the unscheduled maintenance occurs. In the 4^{th} quarter, a scheduled maintenance occurs on unit 5, and an unscheduled maintenance on unit 1 in the 12^{th} week. In the 13^{th} week, only one unscheduled maintenance takes place for unit 1. Figure 108 shows how the power plant will operate when this unscheduled maintenance occurs. In the 12^{th} week, because more than one maintenance activity is planned, generating customer demand becomes more difficult because of the limited system generation capacity. The online generation units have to switch their operating conditions to peak load to remedy the loss of generation. In addition, performing maintenance is also more challenging because of the limited maintenance resources.

Table 35: Deviation 3: Unscheduled Maintenance

Quarter	Week	Unit	Duration (Week)
4	12	1	2
ONE CONFLICT			

Table 36 shows all the maintenance activities in the 4^{th} quarter. Figure 109 shows the system generation if no action is taken during the maintenance window and if system status is switched to a temporary one by adjusting the operating conditions of the generation units. If no action is taken, the system has difficulty satisfying customer demand due to the maintenance. By switching the load level of the other generation units according to Table 37, the system is able to generate enough power to meet customer demand.

Table 38 gives the maintenance activities in the 14^{th} quarter, when only one scheduled maintenance activity takes place compared to the two in the baseline operation. The scheduled maintenance for unit 1 in the second week in the baseline operation is not needed any more. Figure 110 compares the system generation under two conditions, and compares those with customer demand. Table 39 shows the load level changes when maintenance activities occur.

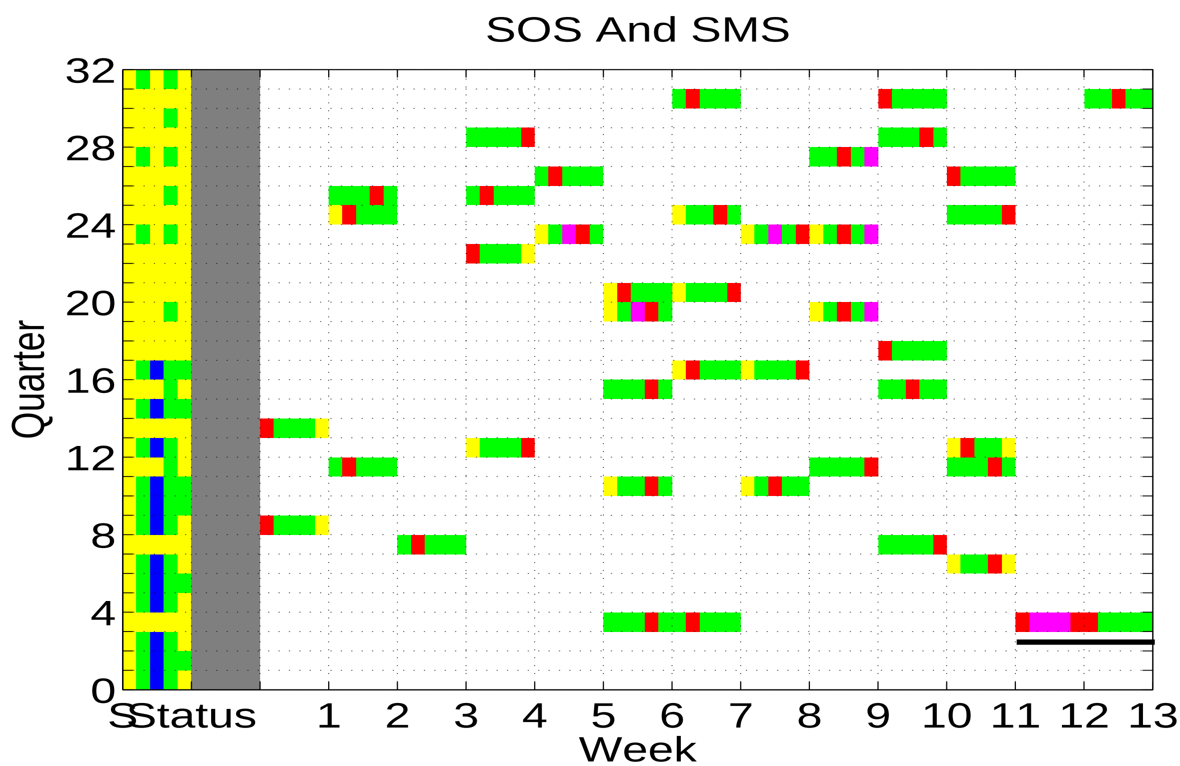


Figure 108: Deviation 3: SOS and SMS

Table 36: Deviation 3: Maintenance Activities in the 4^{th} Quarter

Week	Unit	Maintenance Type
6	4	S
7	2	S
12	5	S
12	1	U
13	1	U

Table 37: Deviation 3: System Status Adjustments in the 14^{th} Quarter

Unit	Before	6^{th} Week	7^{th} Week	12^{th} Week	13^{th} Week
1	PartLoad	BaseLoad	BaseLoad	Maintenance	Maintenance
2	PartLoad	BaseLoad	Maintenance	PeakLoad	BaseLoad
3	PartLoad	BaseLoad	BaseLoad	PeakLoad	BaseLoad
4	PartLoad	Maintenance	BaseLoad	PeakLoad	BaseLoad
5	PartLoad	BaseLoad	BaseLoad	Maintenance	BaseLoad

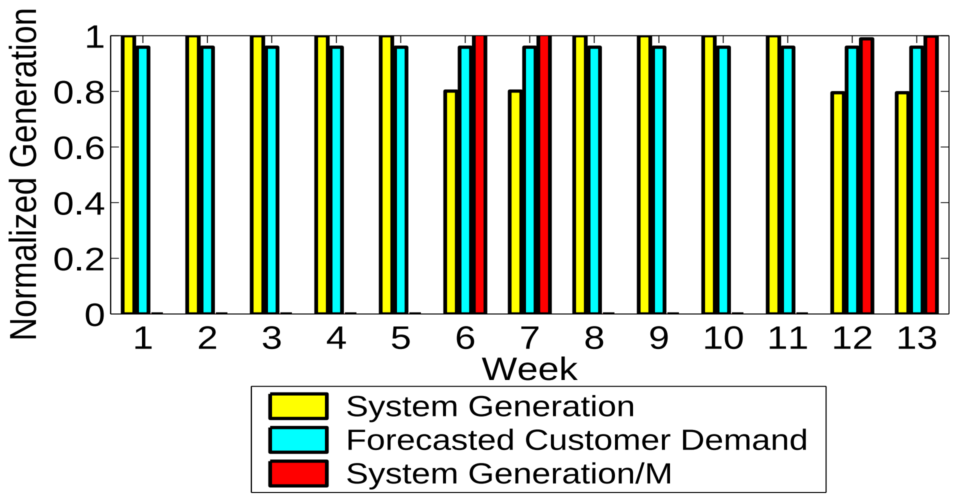


Figure 109: Deviation 3: System Reactions in the 4^{th} Quarter

Table 38: Deviation 3: Maintenance Activities in the 14^{th} Quarter

Week	Unit	Maintenance Type
1	1	S

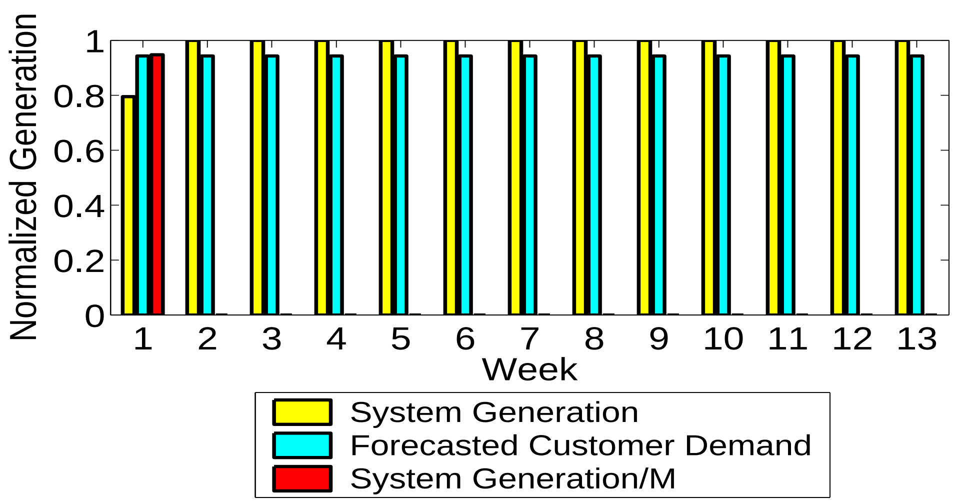


Figure 110: Deviation 3: System Reactions in the 14^{th} Quarter

Figure 111 shows the distributions of fuel costs, maintenance costs, and total costs over the EOP. Total costs peak during the 4^{th} quarter due to unscheduled maintenance. Unscheduled maintenance and its effect on scheduled maintenance cause total costs to rise quickly.

Table 39: Deviation 3: System Status Adjustments in the 14^{th} Quarter

Unit	Before	1^{st}Week
1	PartLoad	Maintenance
2	PartLoad	BaseLoad
3	PartLoad	BaseLoad
4	PartLoad	BaseLoad
5	PartLoad	PartLoad

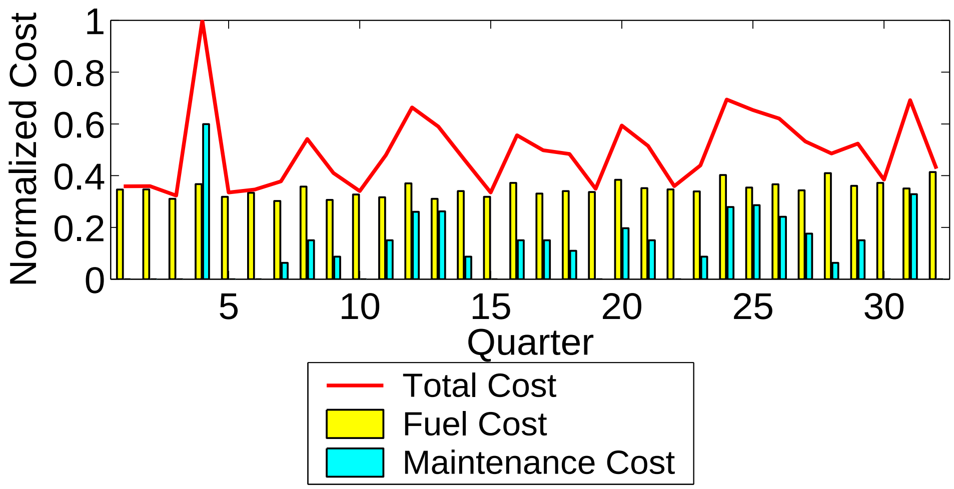


Figure 111: Deviation 3: Power Plant Cost Distributions

5.2.2.4 Deviation 4

An unscheduled maintenance is illustrated in Table 40. In this case, in the 6^{th} week, one unscheduled maintenance occurs on unit 1, and one scheduled maintenance is planned on unit 4. In the 7^{th} week, one unscheduled maintenance occurs on unit 1 and one scheduled maintenance for unit 2. Figure 112 shows how the power plant will operate when this unscheduled maintenance takes place.

Table 40: Deviation 4: Unscheduled Maintenance

Quarter	Week	Unit	Duration (Week)
4	6	1	2
TWO CONFLICTS			

Table 41 illustrates the maintenance activities in the 4^{th} quarter. Figure 113 shows the

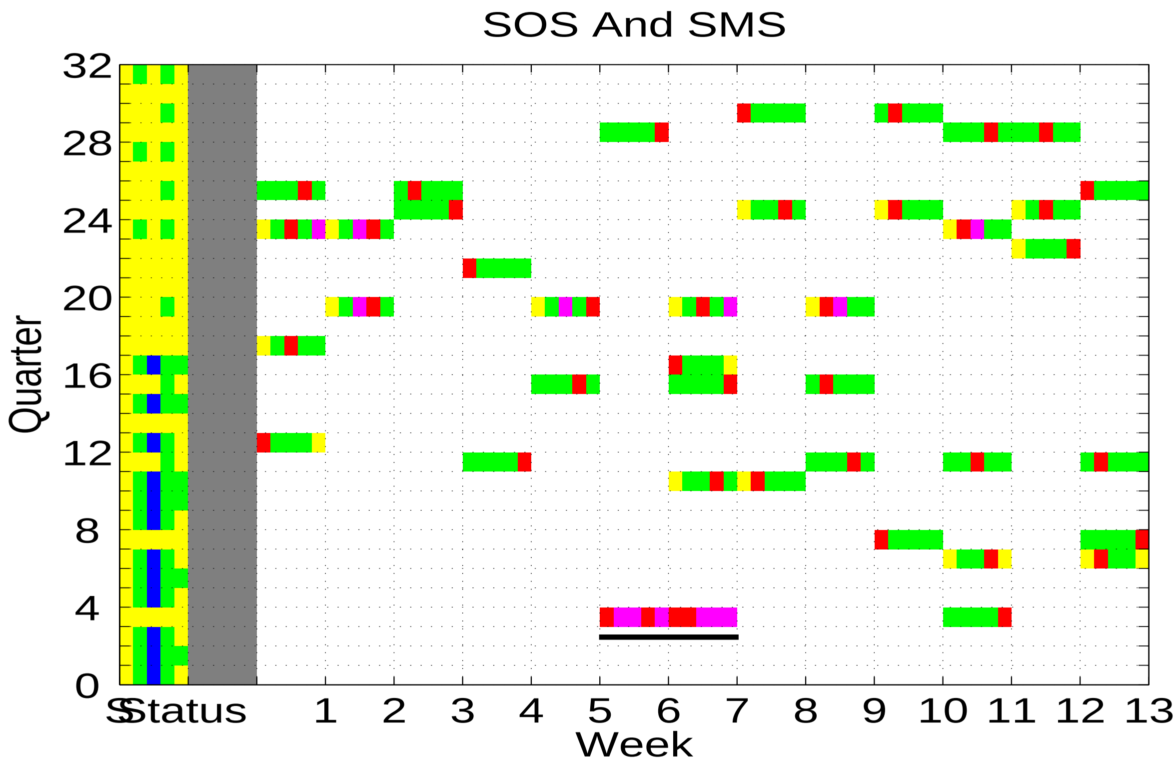


Figure 112: Deviation 4: SOS and SMS

system generation under different conditions and customer demand. If no action is taken, the system can not meet customer demand because of the generation units that are taken offline for maintenance. By switching the load levels of other generation units according to Table 42, the system can achieve a status at which it is able to meet customer demand.

Table 41: Deviation 4: Maintenance Activities in the 4^{th} Quarter

Week	Unit	Maintenance Type
6	4	S
6	1	U
7	2	S
7	1	U
11	5	S

No maintenance activities occur in the 14^{th} quarter. Due to the introduction of maintenance activities at an early time, the recommended maintenance schedules in the baseline operation can not be followed. The maintenance activities have been shifted to the early

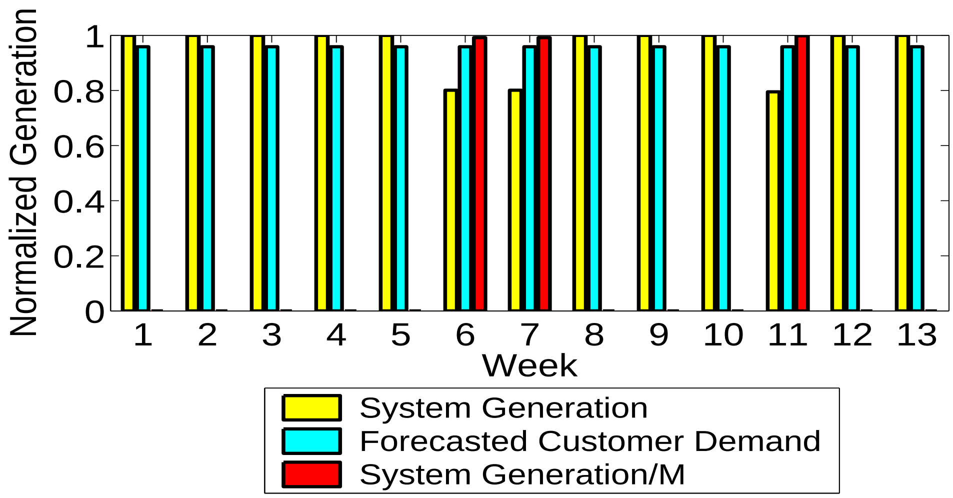


Figure 113: Deviation 4: System Reactions in the 4^{th} Quarter

Table 42: Deviation 4: System Status Adjustments in the 4^{th} Quarter

Unit	Before	6^{th} Week	7^{th} Week	11^{th} Week
1	PartLoad	Maintenance	Maintenance	BaseLoad
2	PartLoad	PeakLoad	Maintenance	BaseLoad
3	PartLoad	PeakLoad	PeakLoad	BaseLoad
4	PartLoad	Maintenance	PeakLoad	BaseLoad
5	PartLoad	PeakLoad	PeakLoad	Maintenance

time, which should happen at some time later according to the accumulative FFH.

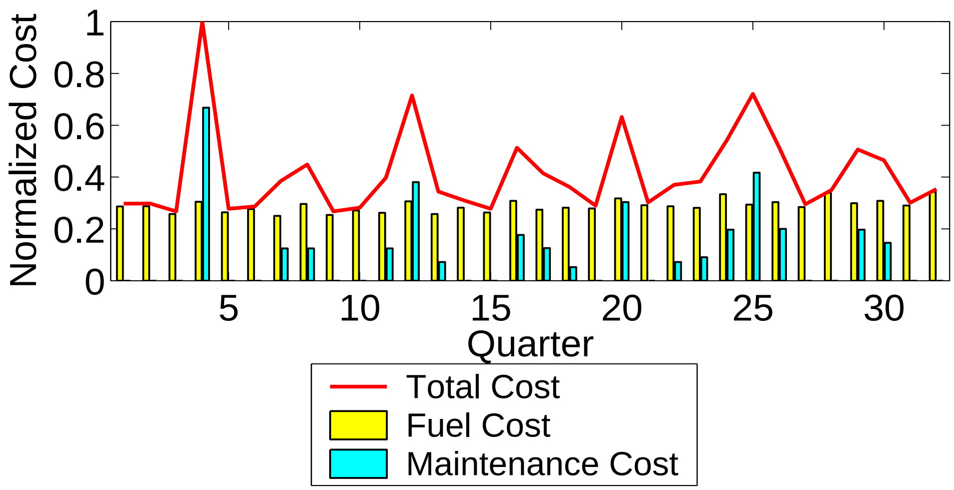


Figure 114: Deviation 4: Power Plant Cost Distributions

Figure 114 shows the distributions of fuel costs, maintenance costs, and total costs. Total costs peak in the 4^{th} quarter during which the unscheduled maintenance is introduced. In this case, in the 6^{th} and 7^{th} weeks, the load levels of other online generation units have been adjusted to their peak load operation. From a demand point of view, an adjustment is made to remedy the huge loss of generation due to two maintenance simultaneously. From an economic point of view, maintenance costs increase due to the difficulty in performing maintenance subjected to limited maintenance resources.

5.2.2.5 Deviation 5

An unscheduled maintenance is introduced at a time such that in the 2^{nd} week of the 14^{th} quarter, both an unscheduled and scheduled maintenance occurs on unit 1. In the 3^{rd} week, only one unscheduled maintenance on unit 1 takes place. Table 43 shows the condition under which the unscheduled maintenance happens. Figure 115 shows how the power plant will operate when this unscheduled maintenance is encountered.

Table 43: Deviation 5: Unscheduled Maintenance

Quarter	Week	Unit	Duration (Week)
14	2	1	2
ONE SAME			

Table 44 illustrates the maintenance activities in the 14^{th} quarter. Figure 116 shows the system generation if no action is taken during the maintenance window and if the system status is switched to a temporary one by adjusting the operating conditions of the generation units. If the load levels of other generation units are switched, the system is able to meet customer demand. The changing of the load level is shown in Table 45. In the 2^{nd} week, despite two maintenance activities, both the unscheduled maintenance and scheduled maintenance occur on same unit, so only one unit is taken offline. Thus, they can actually be treated as one maintenance activity from the generation point of view. Table 45 shows that the load levels of other online generation units are at the base load. Peaking load is not needed because the power plant only loses generation from one unit.

Figure 117 shows the distributions of fuel costs, maintenance costs, and total costs. Total costs peak at the 14^{th} quarter due to the introduction of the unscheduled events.

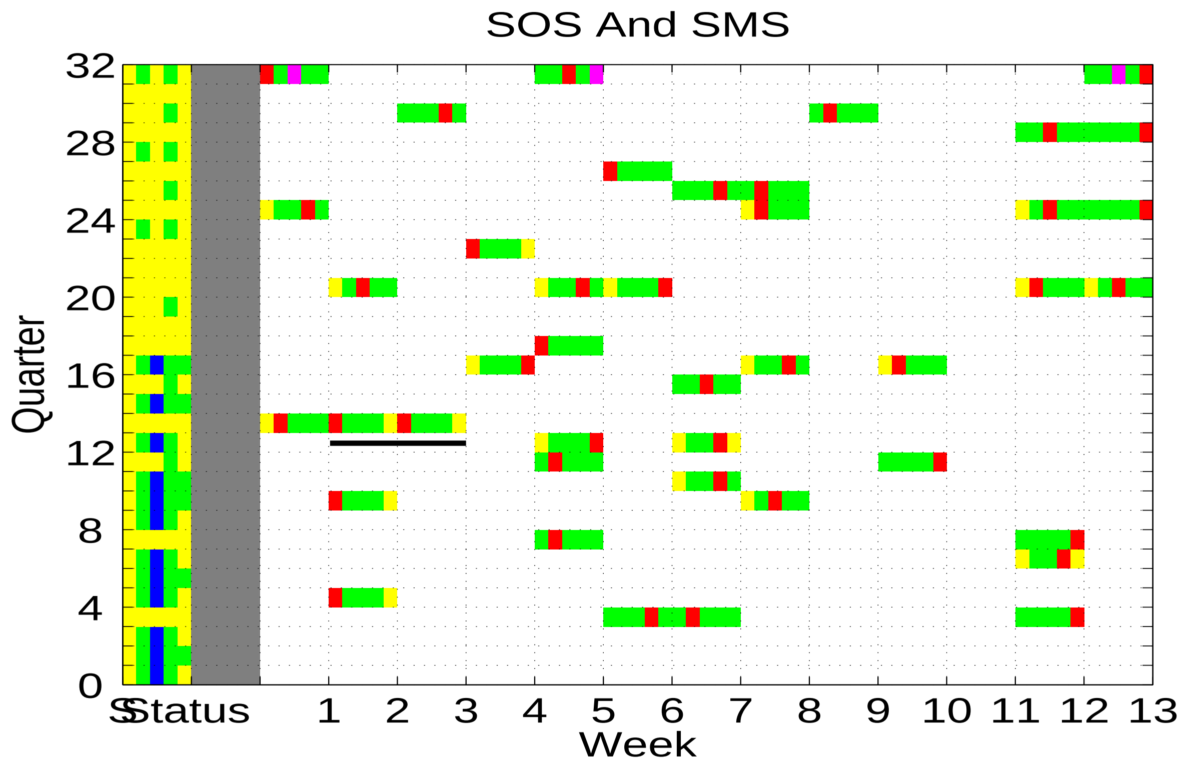


Figure 115: Deviation 5: SOS and SMS

Table 44: Deviation 5: Maintenance Activities in the 14^{th} Quarter

Week	Unit	Maintenance Type
1	2	S
2	1	S&U
3	1	U

5.2.2.6 Deviation 6

In this case, an unscheduled maintenance occurs in the 14^{th} quarter. In the 1^{st} week, a unscheduled maintenance occurs on unit 1 and a scheduled maintenance occurs on unit 2. In the 2^{nd} week, both an unscheduled maintenance and a scheduled maintenance take place on unit 1. Table 46 shows the condition under which the unscheduled maintenance occurs and Figure 118 shows its effect on the operation of the power plant.

Table 47 illustrates the maintenance activities in the 14^{th} quarter. Figure 119 shows the system generation if no action is taken during the maintenance window and if the

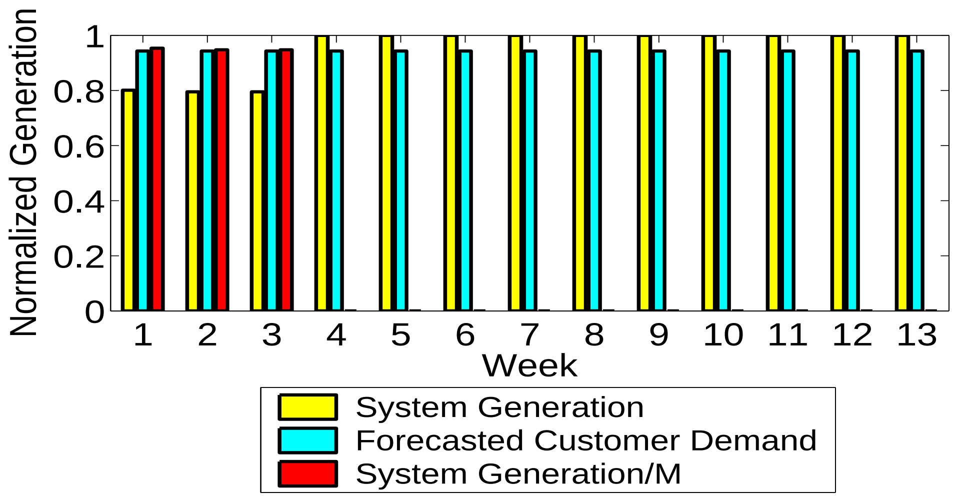


Figure 116: Deviation 5: System Reactions in the 14^{th} Quarter

Table 45: Deviation 5: System Status Adjustments in the 14^{th} Quarter

Unit	Before	1^{st} Week	2^{nd} Week	3^{rd} Week
1	PartLoad	PartLoad	Maintenance	Maintenance
2	PartLoad	Maintenance	BaseLoad	BaseLoad
3	PartLoad	BaseLoad	BaseLoad	BaseLoad
4	PartLoad	BaseLoad	BaseLoad	BaseLoad
5	PartLoad	BaseLoad	PartLoad	PartLoad

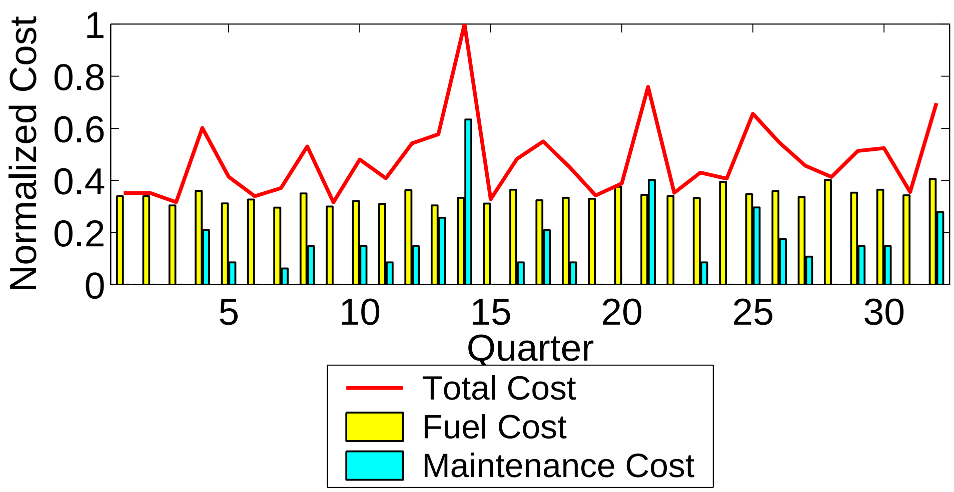


Figure 117: Deviation 5: Power Plant Cost Distributions

system status is switched to a temporary one by adjusting the operating conditions of the generation units. By switching the load level of other generation units according to Table 48,

Table 46: Deviation 6: Unscheduled Maintenance

Quarter	Week	Unit	Duration (Week)
14	1	1	2
ONE SAME, ONE CONFLICT			

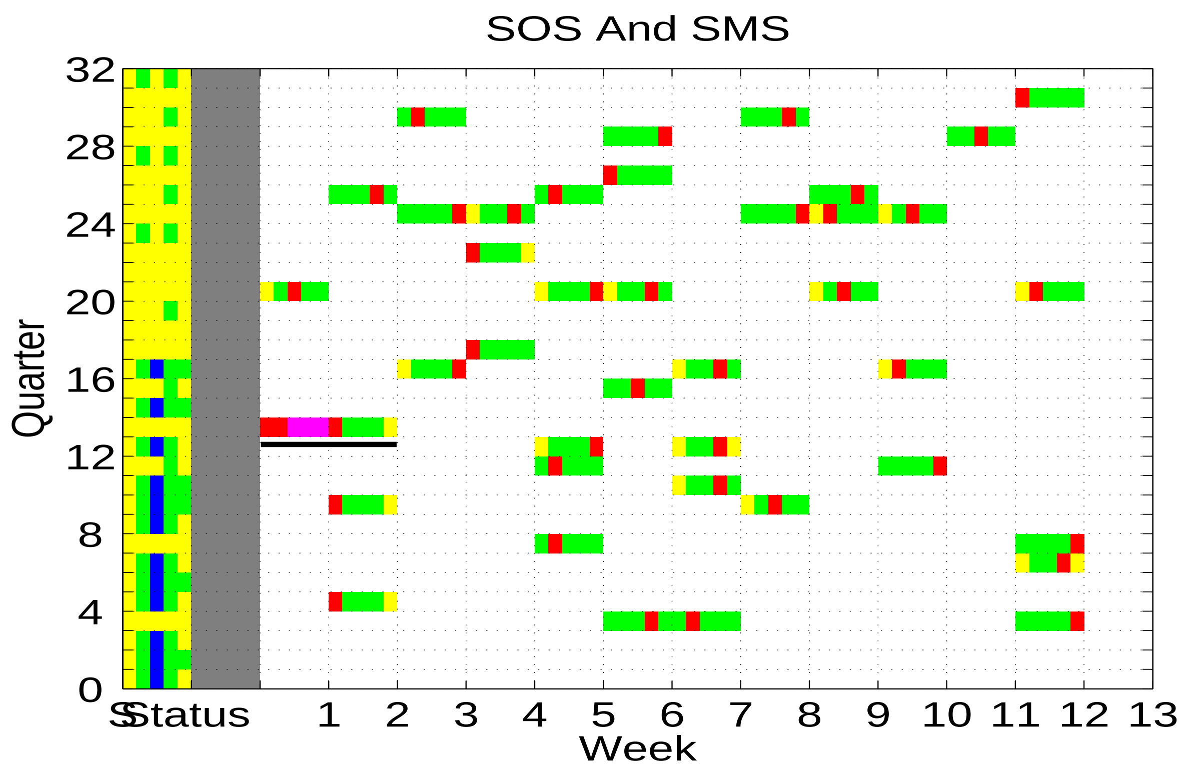


Figure 118: Deviation 6: SOS and SMS

the system is able to meet customer demand. The operating condition in the 1^{st} week is more severe than that in the 2^{nd} week because the power plant has to handle both scheduled maintenance and unscheduled maintenance simultaneously.

Table 47: Deviation 6: Maintenance Activities in the 14^{th} Quarter

Week	Unit	Maintenance Type
1	2	S
1	1	U
2	1	U

Figure 120 shows the distributions of fuel costs, maintenance costs, and total costs over

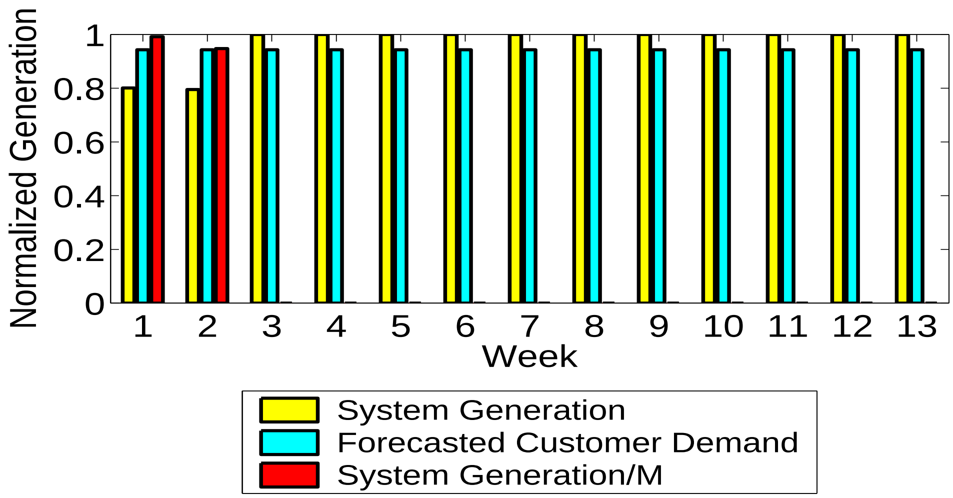


Figure 119: Deviation 6: System Reactions in the 14^{th} Quarter

Table 48: Deviation 6: System Status Adjustments in the 14^{th} Quarter

Unit	Before	1^{st} Week	2^{nd} Week
1	PartLoad	Maintenance	Maintenance
2	PartLoad	Maintenance	BaseLoad
3	PartLoad	PeakLoad	BaseLoad
4	PartLoad	PeakLoad	BaseLoad
5	PartLoad	PeakLoad	PartLoad

the EOP. Total costs again peak in the 14^{th} quarter due to the introduction of unscheduled maintenance.

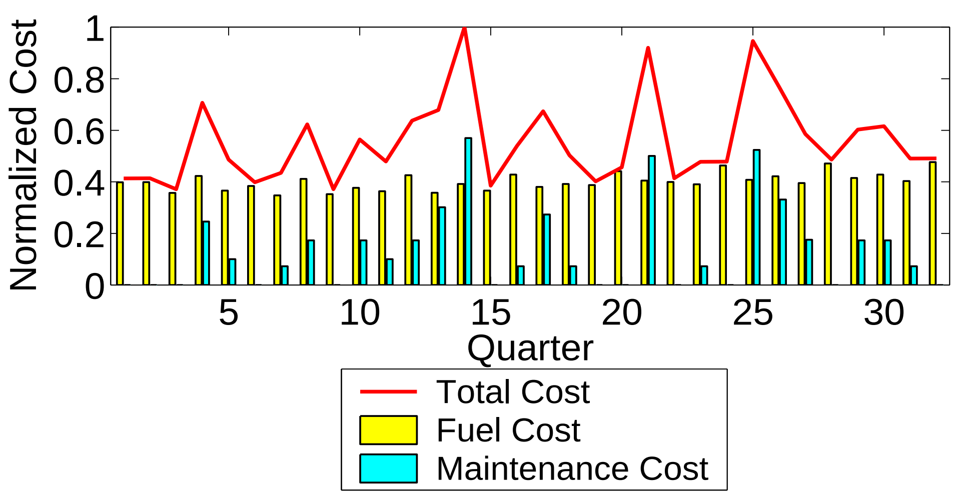


Figure 120: Deviation 6: Power Plant Cost Distributions

Figure 121 shows the total costs associated with the baseline operation and each deviation. If the recommended maintenance schedule can be followed with no unscheduled events interrupting the operation process, the incurred LCC is minimal. The occurrence of unscheduled maintenance activities, however, increases the system LCC and thus diverts the system from the optimal condition, the worst case being simultaneous unscheduled and scheduled maintenance, which contributes to a huge loss of system generation and competion for limited maintenance resources. The LCC increases due to customer demand challenge, lower system reliability, and higher maintenance costs.

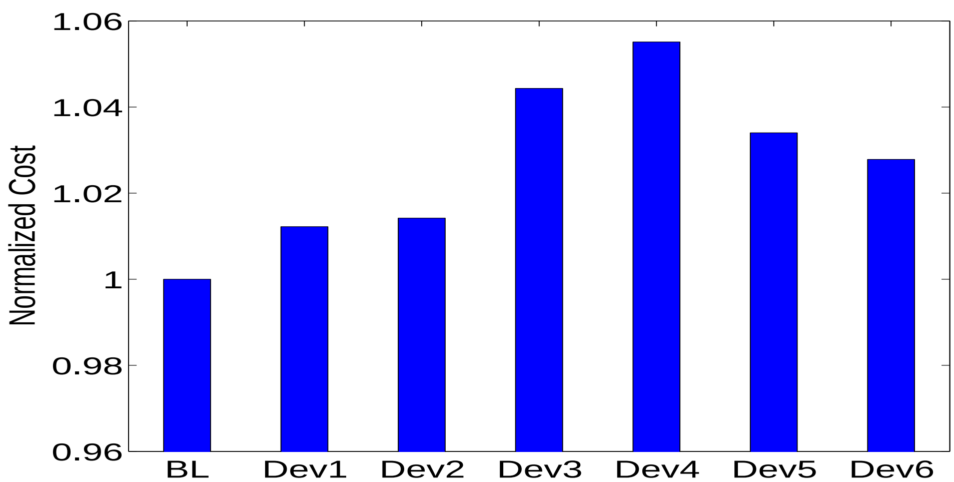


Figure 121: System Total Cost Comparison

5.3 System Capacity Expansion Plans

When the power plant lacks long-term production capabilities, it requires an expansion of system capacity. The EOT =2016, designed for such a capacity expansion, determines the number of generation units needed. According to customer demand forecasting, only one unit is introduced into the power plant. Table 49 shows the distribution of generation for the power plant.

The rough estimate of the EOP is shown in Figure 122. The EOP and EOT are identified to be

$$\text{EOP} = 20 \text{ Quarters, and}$$

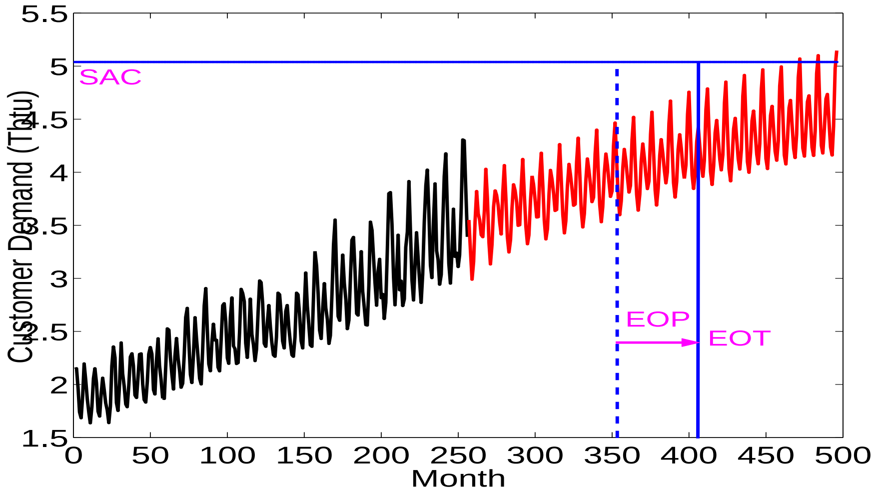


Figure 122: Expansion: Economical Operating Period

$$\text{EOT} = \text{Nov. 2016 Yr.}$$

Now the DM process for the expanded power plant can be carried out in the same way as that for the baseline system. Figure 123 shows how the system will operate during capacity expansion. The far left column shows the system status for each quarter. Beyond 32 quarters, expansion has been carried out and the system has 6 generation units. The figure shows that the new generation unit is not in service when the system is in normal operation, but when any generation units are taken offline for maintenance, scheduled or unscheduled, the new generation unit is needed, as it helps the system satisfy customer demand during generation contingencies. As customer demand increases, as forecasted, the system will utilize the new generation unit during the normal operation.

Table 49: Expansion: Normalized Generation Unit Output

Unit	Part Load	**Base Load**	Peak Load	Maintenance	Off
1	0.6137	**0.7659**	1.00	0	0
2	0.5962	**0.7484**	0.9912	0	0
3	0.5787	**0.7309**	0.9825	0	0
4	0.5962	**0.7484**	0.9912	0	0
5	0.6137	**0.7659**	1.00	0	0
6	0.5962	**0.7309**	0.9737	0	0
System Capacity $= \mathbf{4.4904}\mathcal{HUP}$					
SAC $= \mathbf{3.5923}\mathcal{HUP}$					

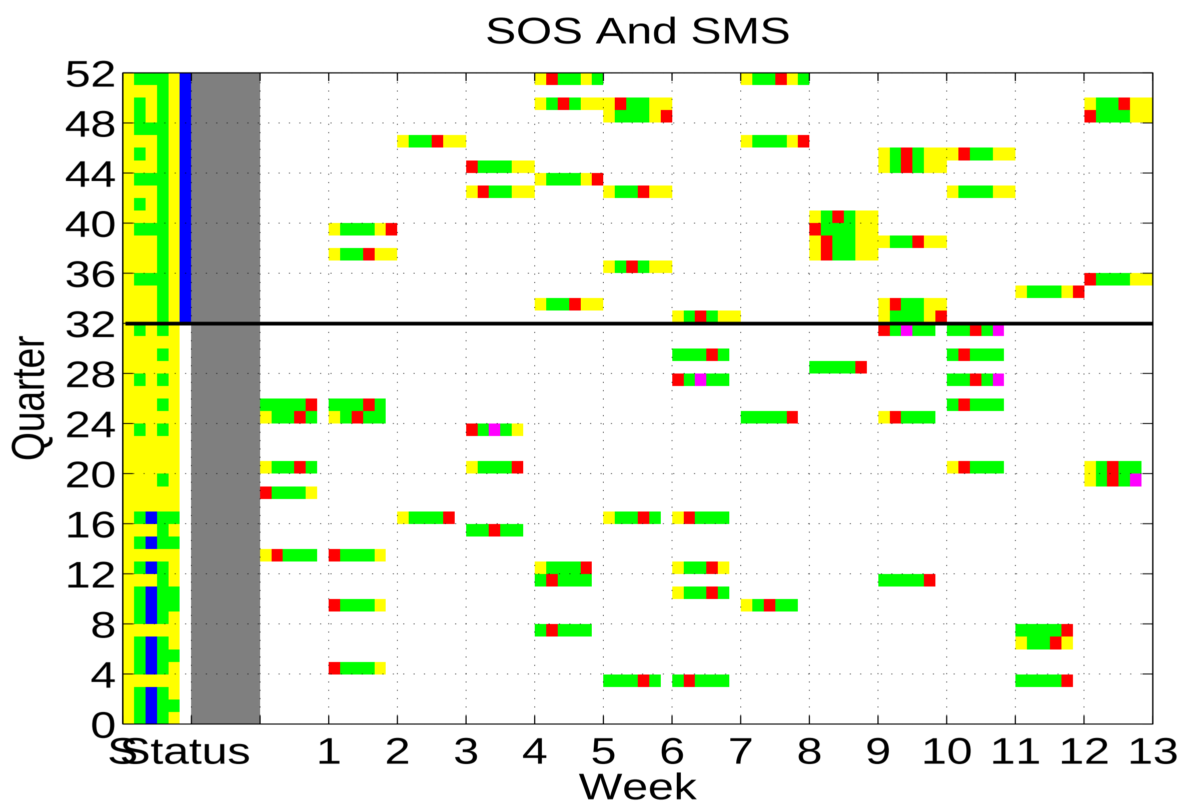


Figure 123: Expansion: SOS and SMS

Figure 124 shows the system output and forecasted customer demand. The power plant under the SOS is able to satisfy customer demand and capture the seasonal variations, too.

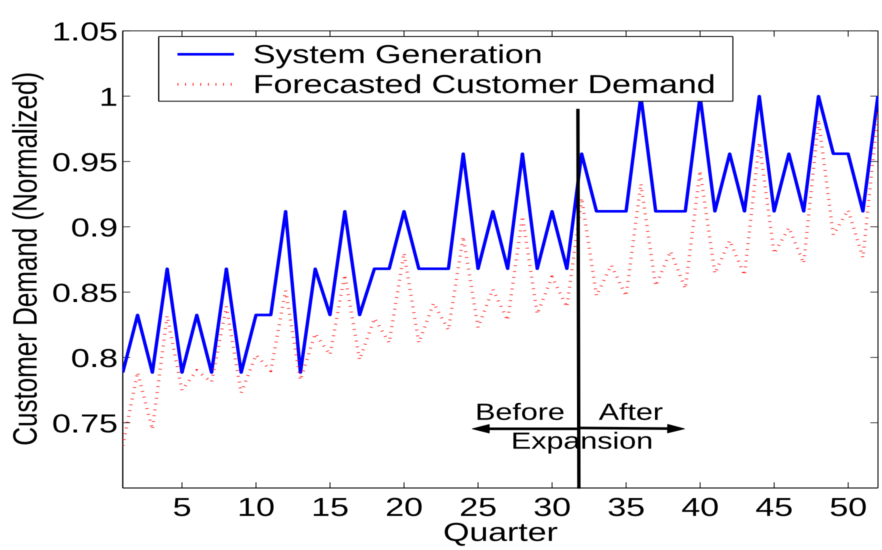


Figure 124: Expansion: System Generation vs. Customer Demand

The cost associated with the expansion of a power plant includes fuel costs, maintenance costs, investment costs, and so forth. The investment costs are levelized over the whole expansion period. Figure 125 shows the distributions of fuel costs, maintenance costs, and total costs over the whole EOP. The total is 8.852NV.

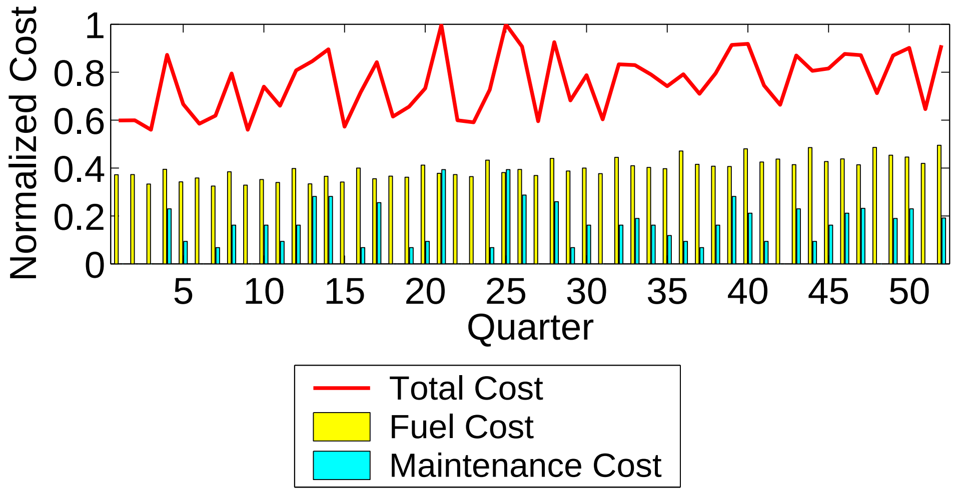


Figure 125: Expansion: System Cost Distributions

5.4 A Bootstrapping Estimate of the LCC

The bootstrap method is used to measure the bias of the estimated system LCC to the actual LCC needed to drive business. Block bootstrapping is performed on the historical data to generate pseudo samples that are utilized as input to the forecasting method WAW. Figures **??**, **??**, and **??** show one pseudo sample for customer demand, natural gas prices, and electricity prices, respectively. There are a total of 20 such samples for each. Block bootstrap generates pseudo samples by keeping the internal structure of the data series.

Each set of forecasting results based on the pseudo samples is used as an input to the DM process. An optimal operating strategy is chosen to achieve the minimal total LCC for each of them. Table 50 gives the total LCC associated with each of these SOS. The average LCC based on these pseudo samples is 3.6685$\mathcal{NV}$ or 0.9963$\mathcal{BLNV}$, where $\mathcal{BLNV}$ is the baseline total cost. The bias is calculated to be the difference between the baseline value and the estimated value, which is -0.0029 $\mathcal{NV}$ or -0.0037 $\mathcal{BLNV}$. This means that the baseline overestimates the LCC that is actually needed to drive business for the power plant on average, but only by a small amount. The histogram is shown in Figure 126. Nearly all the estimates of the LCC lie in the interval [3.3$\mathcal{NV}$, 4.7$\mathcal{NV}$].

Table 50: LCC for Each Pseudo Sample

Sample	1	2	3	4	5	6	7	8	9	10
LCC($\mathcal{NV}$)	4.0122	4.1625	3.6104	3.7138	3.7755	3.6730	3.9815	3.7767	3.5305	3.5854
Sample	11	12	13	14	15	16	17	18	19	20
LCC($\mathcal{NV}$)	3.2859	3.5966	3.8379	3.4798	4.5126	4.4262	3.6841	4.7248	3.4489	3.3848

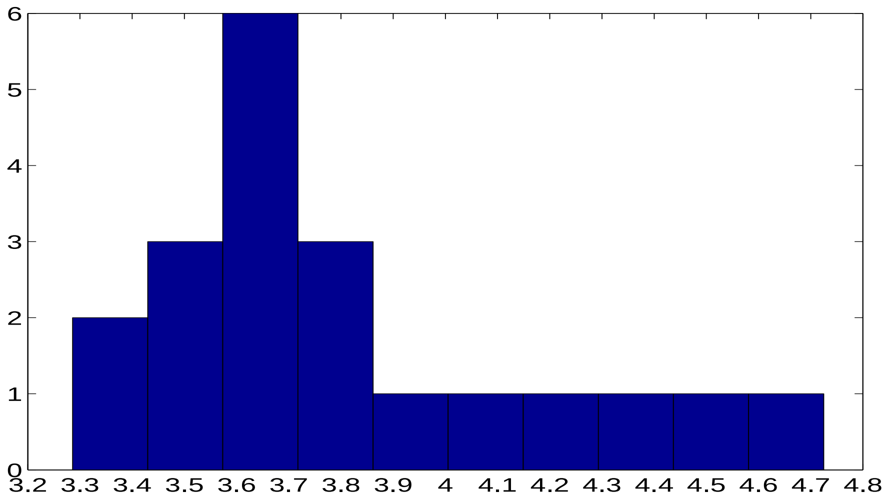


Figure 126: Histogram of Total LCC

5.5 Uncertainty Exploration

Two factors, weather and economic development, are known as very important driving forces in the electric market [33]. Two indicators that represent their functions have been chosen. For each factor, a vector is used to describe its condition. For factor W, the value varies from -1$\mathcal{NI}$ to 1$\mathcal{NI}$, where $\mathcal{NI}$ is a normalized value. The second element represents the time that an external force occurs, which varies from the 4^{th} month to the 34^{th} month of the EOP. The third element represents the duration of the impact of this factor, whose value is fixed at 3 months. For factor E, the value varies from -1$\mathcal{NI}$ to 1$\mathcal{NI}$. The second element varies from the 4^{th} month to the 34^{th} month. The third element has a fixed value of 12 months. The time lag is assumed to be 3 months from the time the phenomenon occurs to the time that it has an impact on the system. The time lag is assumed to be fixed because the impact of varying it is the same as the varying of the time at which it occurs. Table 51 shows the morphological fields for these two factors separately.

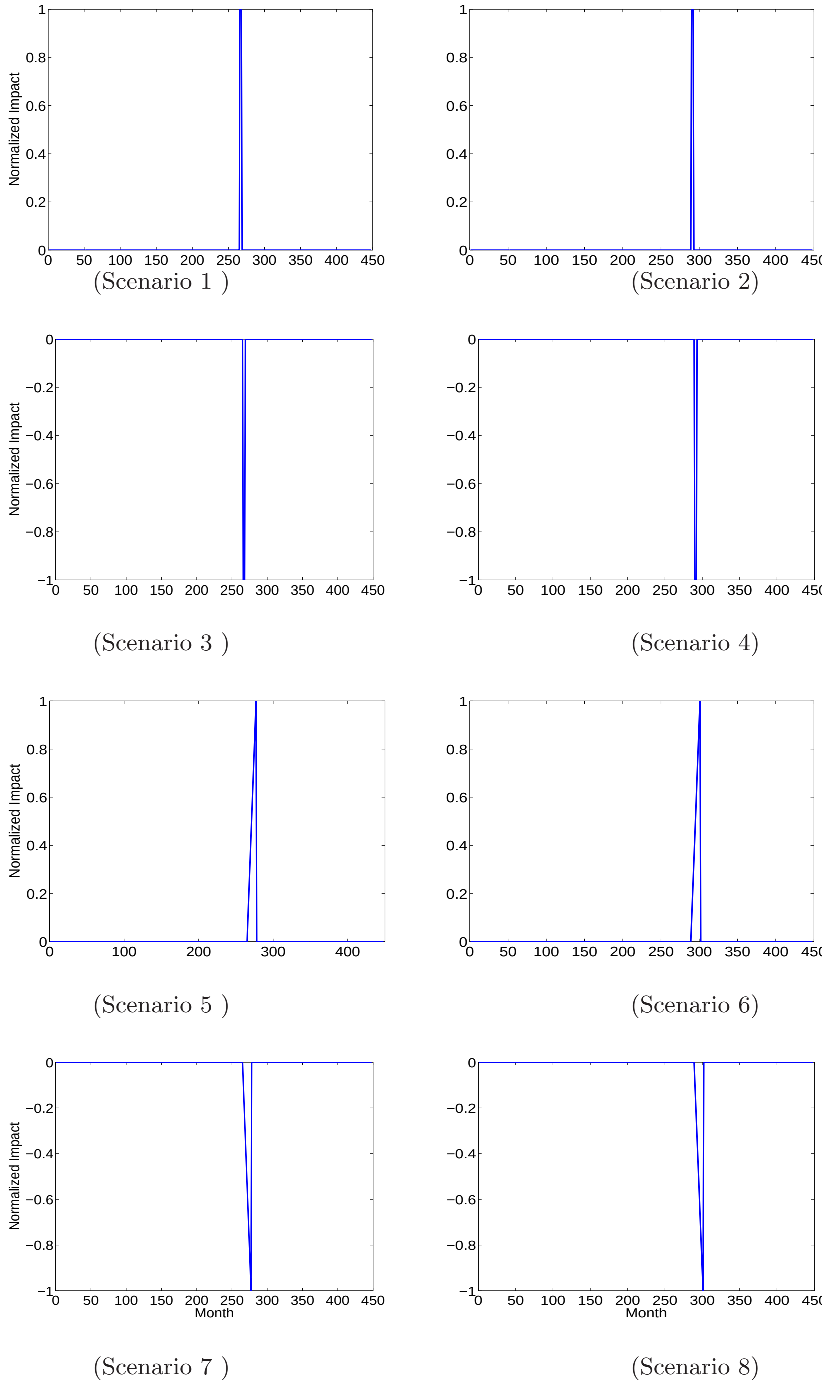
Table 51: Morphological Fields For Parameters

W Factor		E Factor	
Value	Time	Value	Time
1 $\mathcal{NI}$	4	1 $\mathcal{NI}$	4
-1 $\mathcal{NI}$	34	-1 $\mathcal{NI}$	34

Figure 127 shows the eight scenarios corresponding to those listed in the above matrix. Factor W is a phenomenon that occurs instantaneously and disappears instantaneously. Factor E occurs gradually and disappears instantaneously. These two formats are utilized in order to simulate the impact of the weather and the economy on the power plant.

Figure 128 shows customer demand that is forecasted under each scenario. Figure 128 (1) and (2) shows the impact of external factor W on the forecasting process. The impact of the external forces is an increase in customer demand. The baseline operation has an EOP of 32 quarters. As illustrated, the increase in customer demand caused by these external driving forces has a direct impact on the EOP. Scenario 1 has an EOP of 28 quarters and scenario 2 has an EOP of 24 quarters. Figure 128 (3) and (4) shows the negative impact from external factor W. Now the EOP is 36 quarters for both scenarios. If factor W causes an increase in customer demand, then the later the introduction time, the larger the impact. However, if it causes a decrease in customer demand, the introduction time does not have an obvious impact.

Figure 128 (5) and (6) shows the impact of external factor E. In scenario 5, the system capacity actually meets customer demand before the EOT. Considering it is a short-term demand contingency, the EOP can be extended to 26 quarters. In scenario 6, when a big spike in customer demand occurs, the average customer demand already exceeds the system capacity. The EOP in this case is determined to be 26 quarters, which is the same as that in scenario 5. Figure 128 (7) and (8) shows the impact of external factor E, but it acts



(Scenario 1) (Scenario 2)

(Scenario 3) (Scenario 4)

(Scenario 5) (Scenario 6)

(Scenario 7) (Scenario 8)

Figure 127: Scenarios

inversely, so it decrease customer demand. Its impact is more obvious than the impact of factor W in scenarios 3 and 4, so the EOP now has be 56 quarters and 52 quarters, respectively. Thus, if factor E causes a decrease in customer demand, the the earlier the introduction time, the larger the impact.

Figure 129 shows the power plant generation and customer demand for each scenario. The perturbation incurred by the introduction of external factors causes variations in customer demand; thus, the SOS must ensure that the power plant can minimize total cost while meeting customer demand. From Figure 129, it can be seen that the generation of the power plant can adapt to each scenario by producing the required customer demand.

Figures 130 to 137 describe how the power plant will operate under each scenario. The SOS is listed on the far left column of each figure. The right part of each figure describes the SMS. The increase in customer demand caused by external factors tends to reduce the EOP. Figures 130, 131, 134, 135 show how the system will operate with an increase in customer demand. Their EOP are 28, 24, 26, and 26 quarters, respectively. The impact of the negative external factors is to decrease the rate at which customer demand increases. Figures 132, 133, 136, 137 show that the EOP increases correspond to that in the baseline operation, but the impact of factors W and E are not the same. Factor E influences the system more clearly that factor W. The EOP has increases to 56 and 52 quarters in the last two scenarios.

The distribution of the total LCC over the EOP for each scenario is shown in Figure 138. Maintenance costs and fuel costs, the two major cost components of the total life cycle, are also shown in the figure. The total life cycle over the EOP is not only related to how the power plant operates but also depends on the duration of the EOP. Table 52 shows the total cost and total cost per quarter for each scenario. It can be seen that regardless of what operating strategy the system adopts and how customer demand will varies, the total cost per quarter, which varies from 0.02926 $\mathcal{BLNV}$ to 0.03399 $\mathcal{BLNV}$, is quite stable. The average value for this metric is 0.03205 $\mathcal{BLNV}$, which varies with a range of -8.7 % to 6.1 %. The baseline operation estimated this value at 0.03125 $\mathcal{BLNV}$, which is also within the range. Therefore, the conditions assigned to each factor are wide enough to cover a reasonable

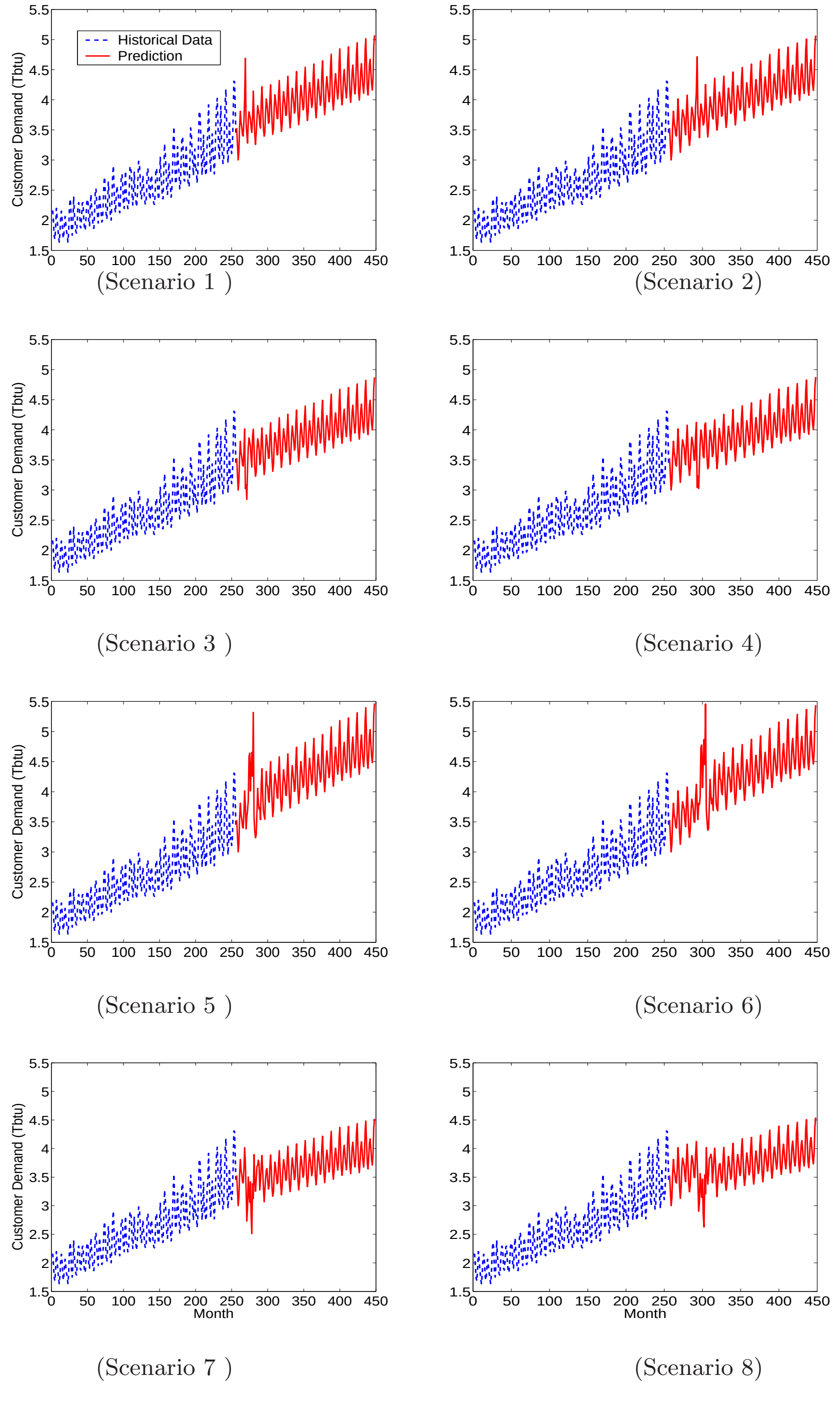


Figure 128: Customer Demand Forecasted Under Each Scenario

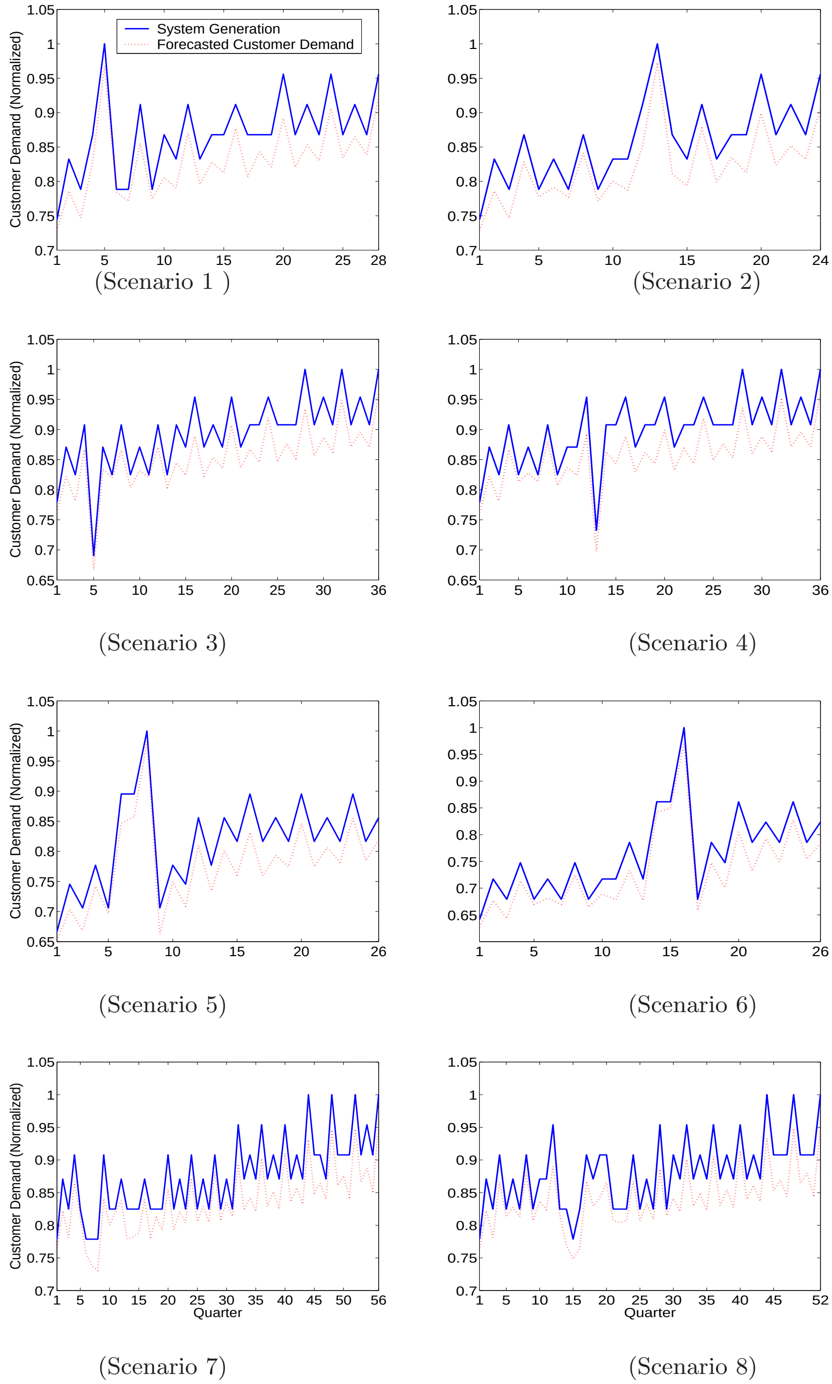


Figure 129: System Generation vs. Customer Demand Under Each Scenario

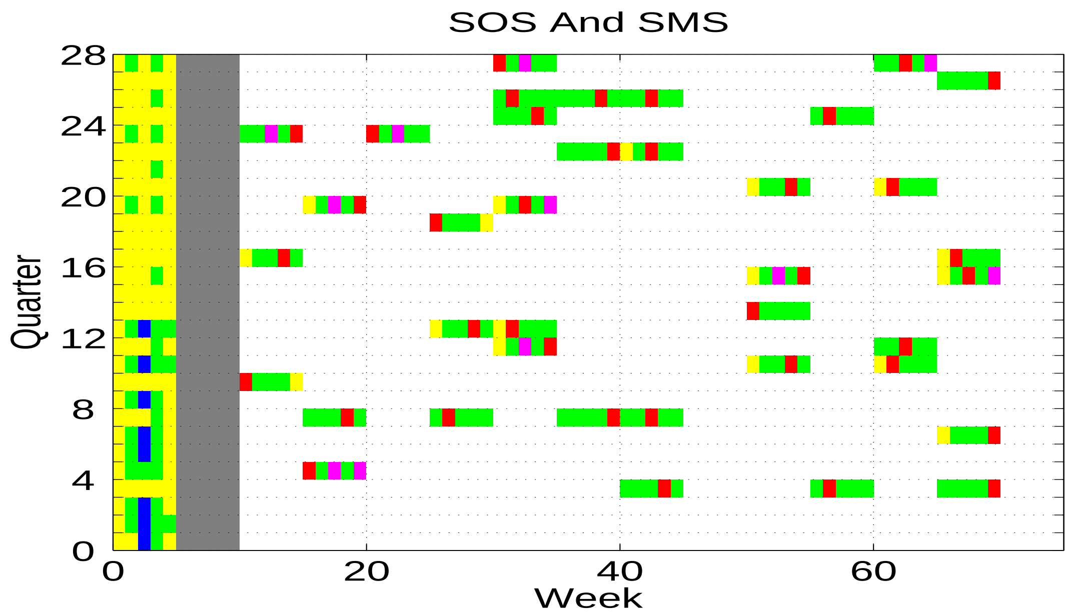


Figure 130: Scenario 1: SOS and SMS

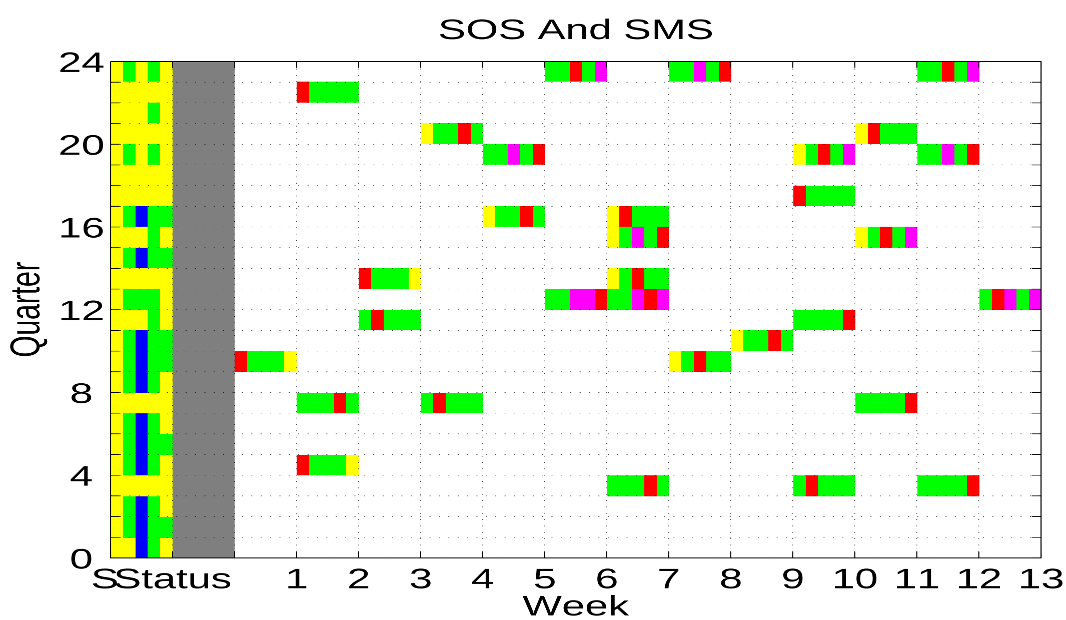


Figure 131: Scenario 2: SOS and SMS

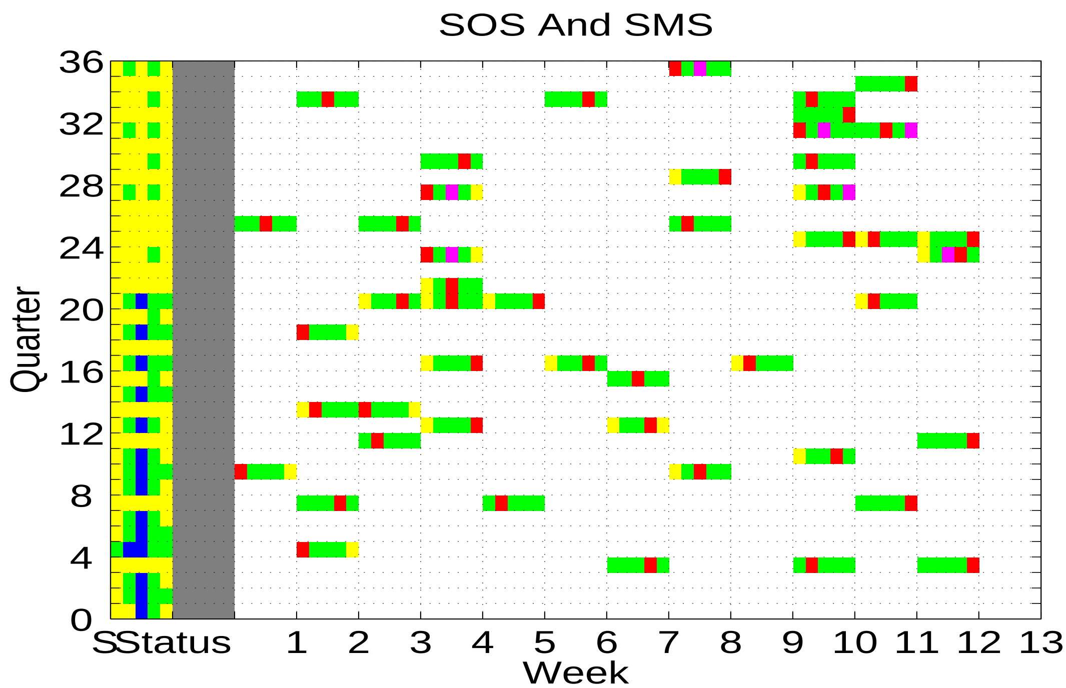


Figure 132: Scenario 3: SOS and SMS

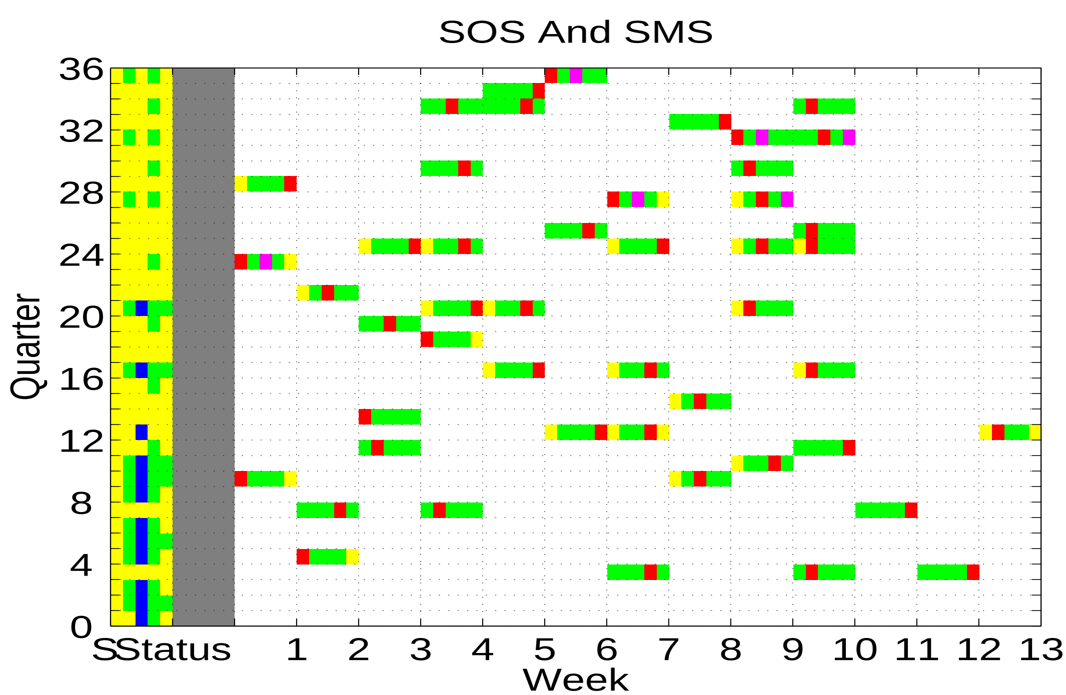


Figure 133: Scenario 4: SOS and SMS

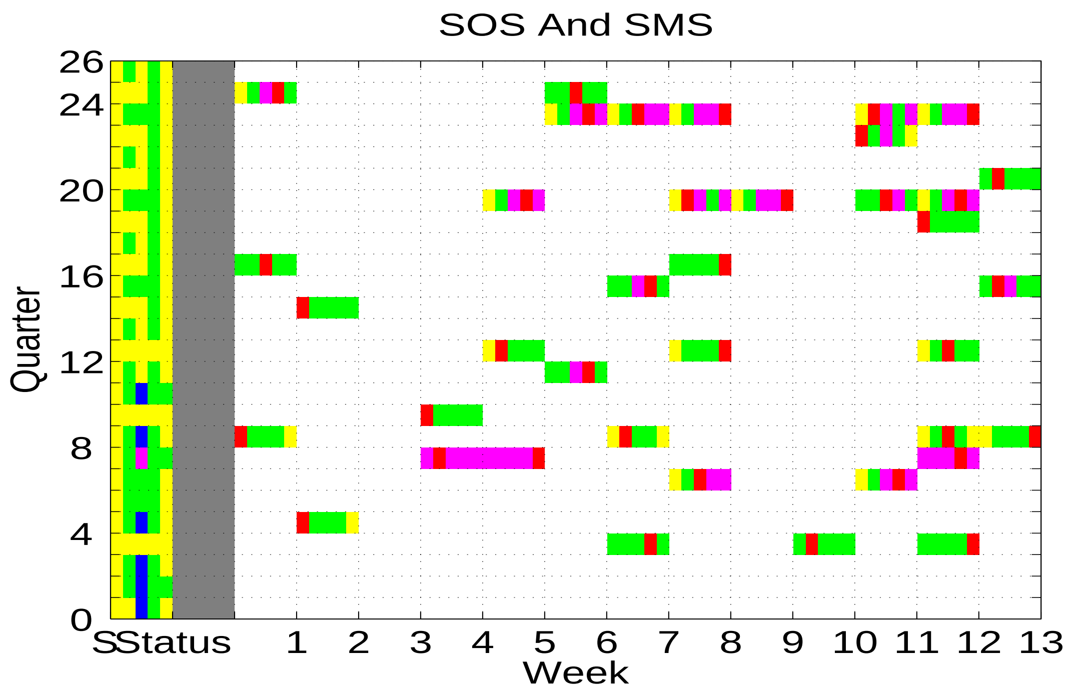


Figure 134: Scenario 5: SOS and SMS

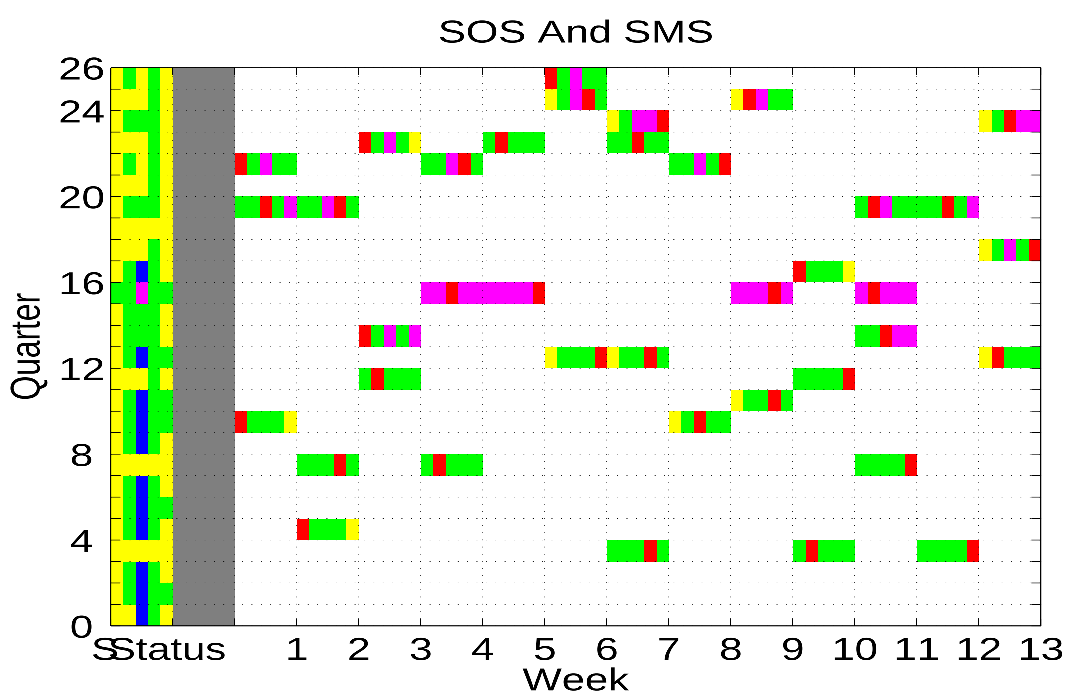


Figure 135: Scenario 6: SOS and SMS

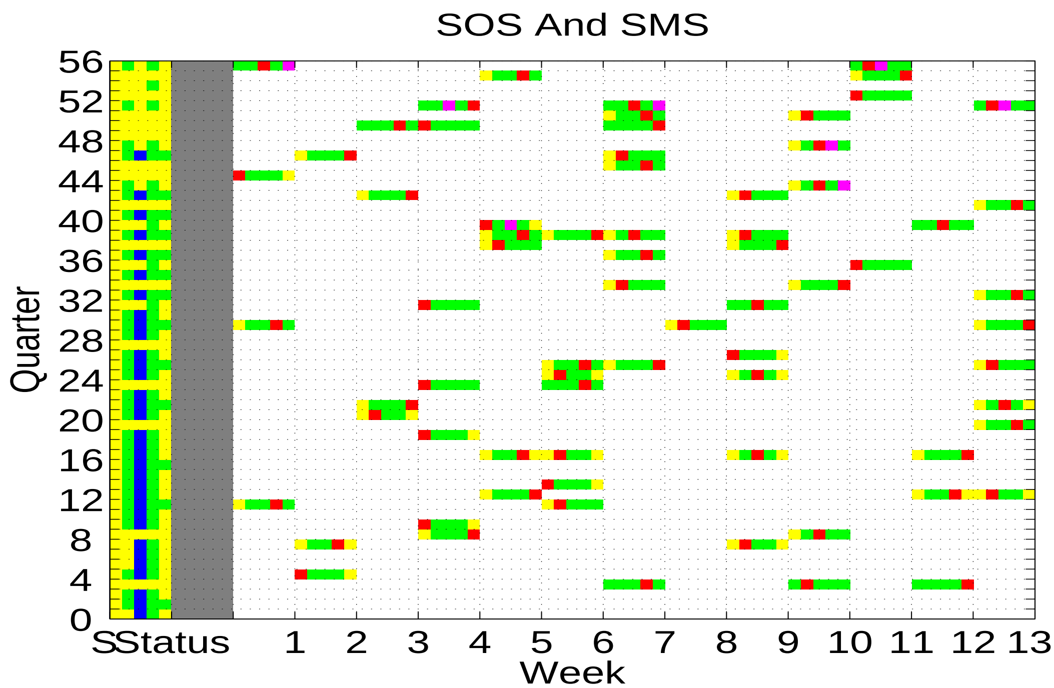


Figure 136: Scenario 7: SOS and SMS

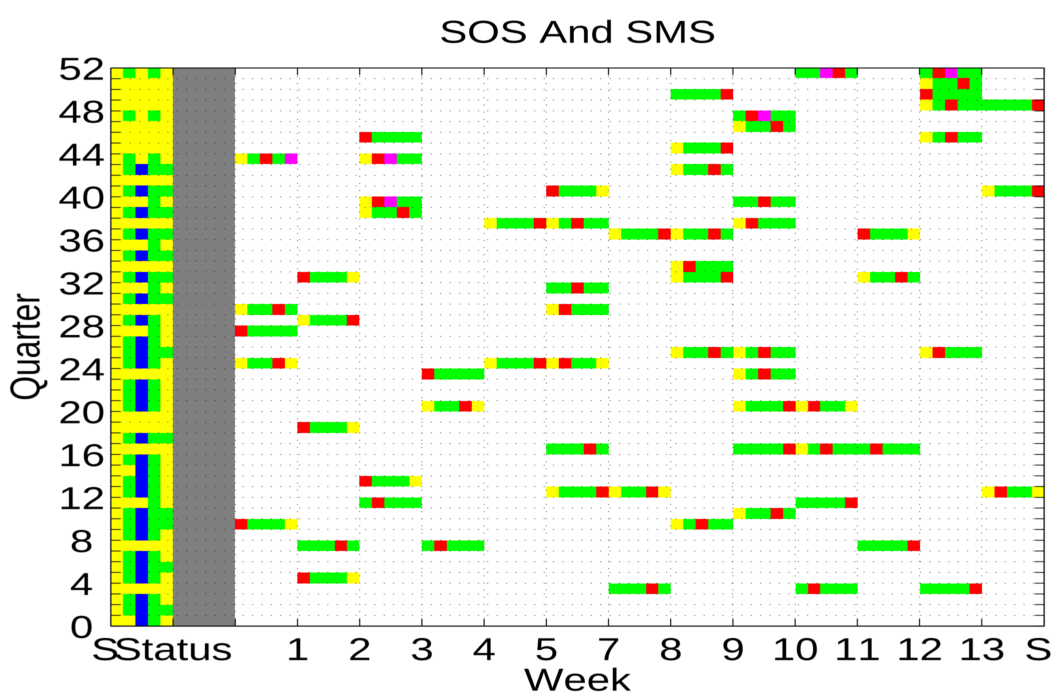


Figure 137: Scenario 8: SOS and SMS

variety of developments and to span the problem space.

Table 52: Total Cost for Each Scenario

Scenario	1	2	3	4	5	6	7	8
Total Cost ($\mathcal{BLNV}$)	0.8980	0.7468	1.1289	1.1276	0.8831	0.8659	1.7674	1.6384
EOP (Q)	28	24	36	36	26	26	52	56
Cost/EOP ($\mathcal{BLNV}$)	0.03207	0.03112	0.03136	0.03132	0.3397	0.03330	0.03399	0.02926

The study of individual factors can clarify the impact of each on customer demand and therefore, on power plant operations. It can also help identify the impact of each factor when more than one external factors act on the system simultaneously. The next step is to combine the morphological fields for these two factors into one morphological field. If considered simultaneously, these two factors render 16 total scenarios. Combinations of these scenarios are shown in Figures 139 and 140. Figures 141 to 142 show customer demand forecasted under each scenario. The forecasting information is input to the DM process to determine the system generation.

Figures 143 and 144 show the production of the system and the customer demand forecasted for each scenario. These figures show that the system is successful in identifying operating strategies that can satisfy customer demand and still capture the seasonal variations in customer demand and the perturbations caused by the introduction of external factors.

The system operation directly impacts the total cost and the cost distributions along the EOP. Figures 145 and 146 show the cost distributions for each scenario. Table 53 gives the EOP for each scenario and the total cost associated with the operation.

The SOS and SMS are shown in Figures 147 to 162. As illustrated, the impact of the second factor on the power plant is much larger than the impact of the first factor. In the first 8 scenarios, the second factor acts as a positive impact on the power plant and regardless of hen the first factor is introduced and whether its impact is positive or negative, the EOP of the power plant shrinks. However, when the second factor acts to decrease customer demand, regardless of when the first factor is introduced and how it impacts on the system,

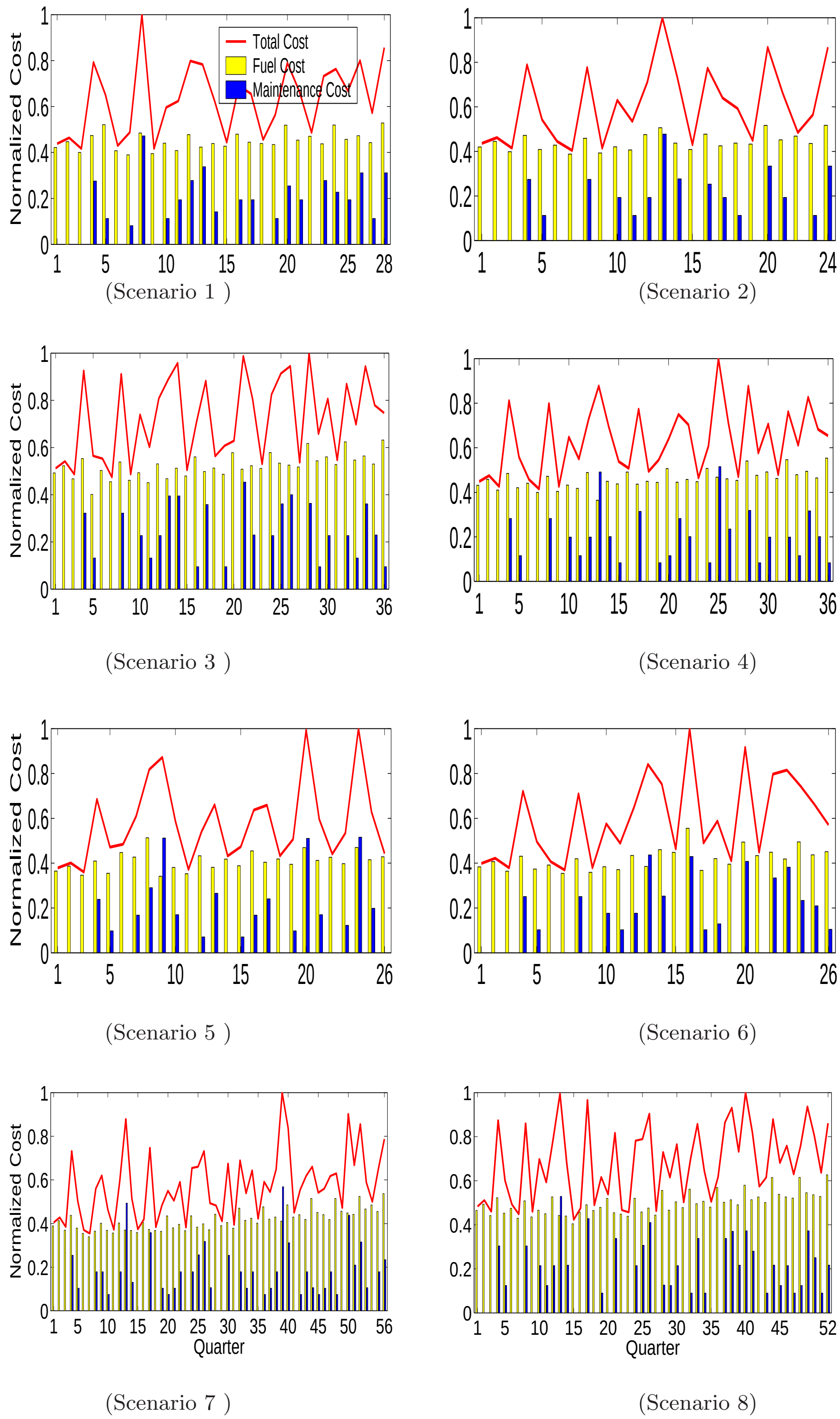


Figure 138: System Total LCC Under Each Scenario

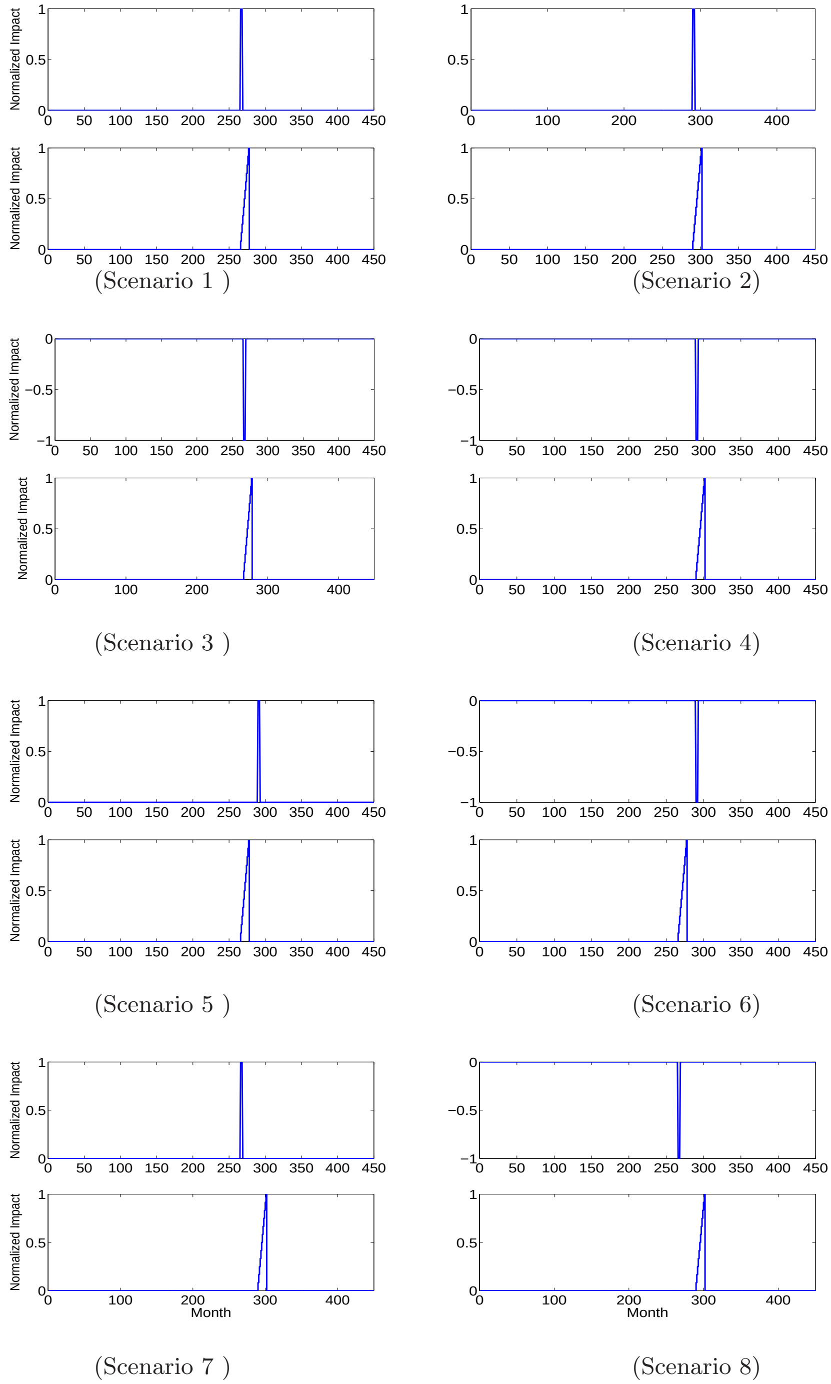
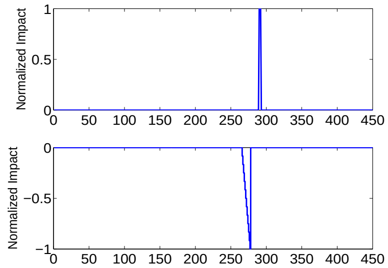
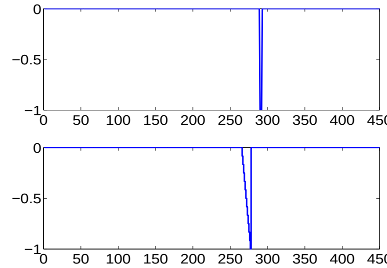


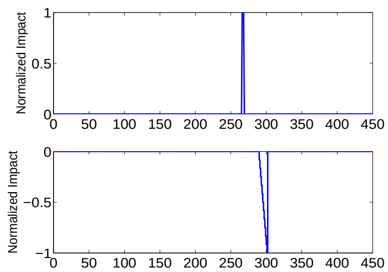
Figure 139: Scenarios (1-8)



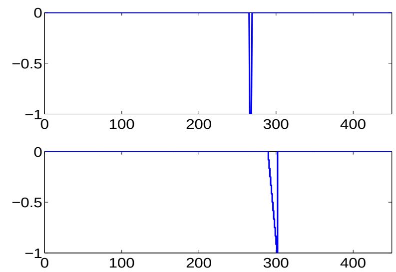
(Scenario 9)



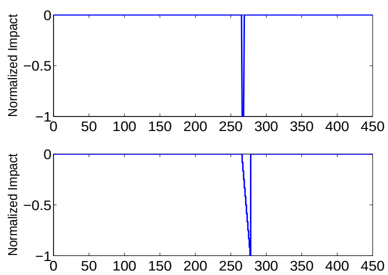
(Scenario 10)



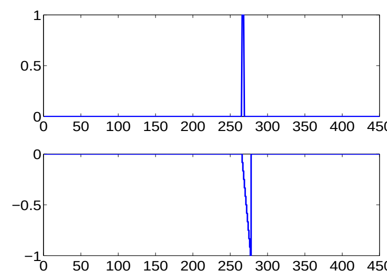
(Scenario 11)



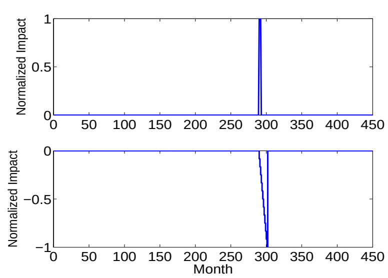
(Scenario 12)



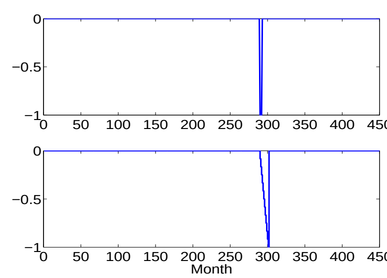
(Scenario 13)



(Scenario 14)



(Scenario 15)



(Scenario 16)

Figure 140: Scenarios (9-16)

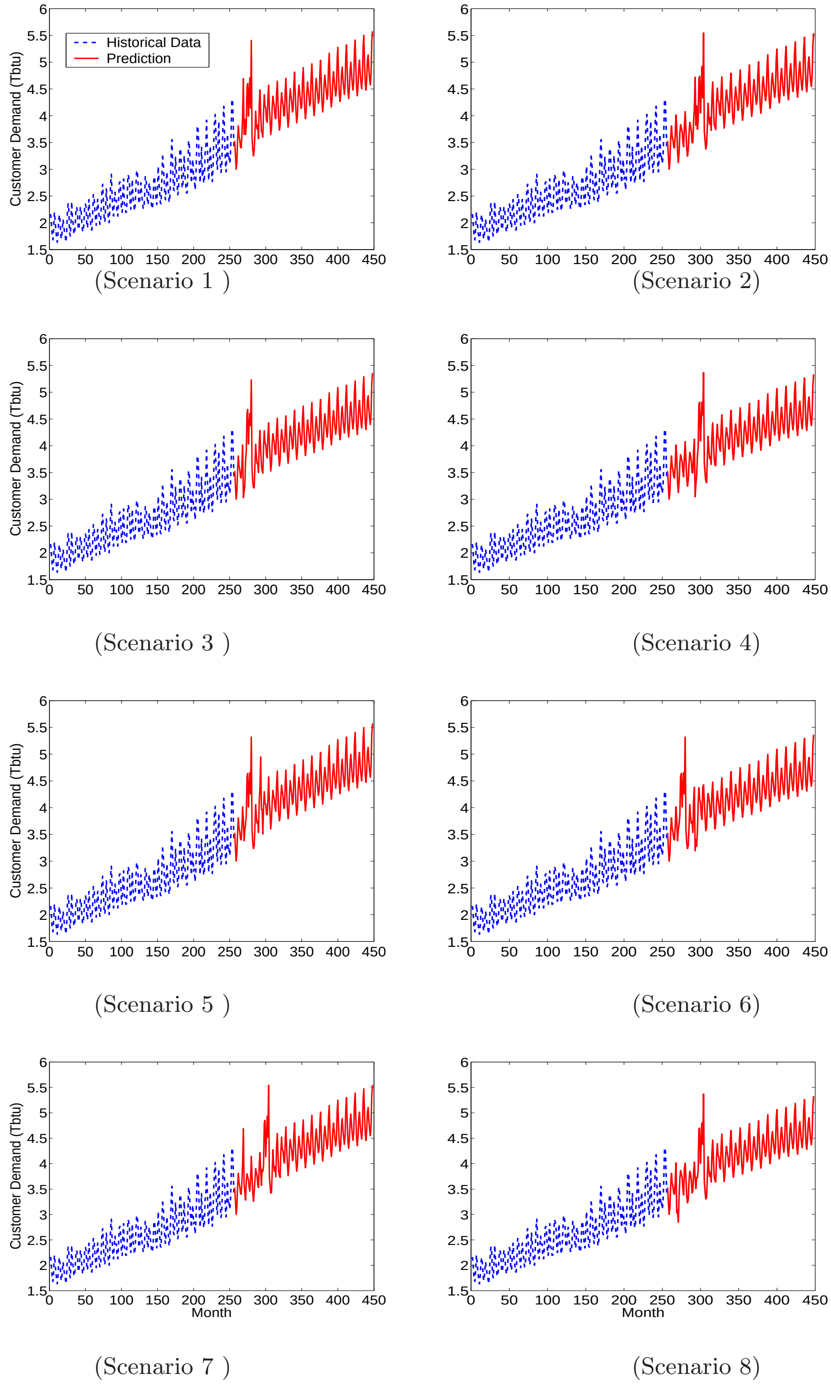
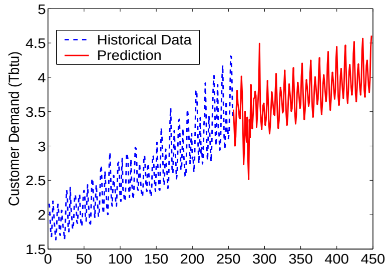
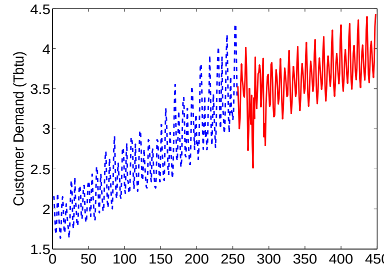


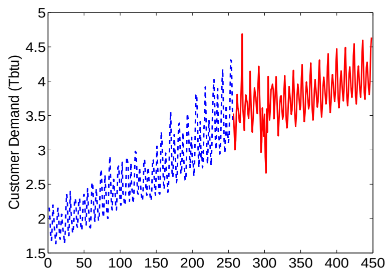
Figure 141: Customer Demand Forecasted Under Each Scenario (1-8)



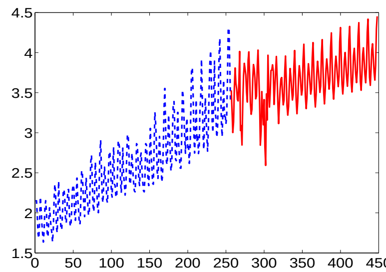
(Scenario 9)



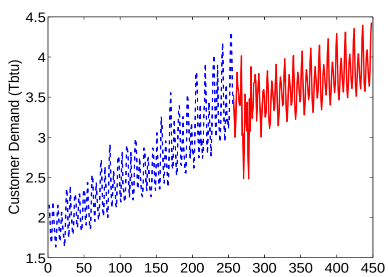
(Scenario 10)



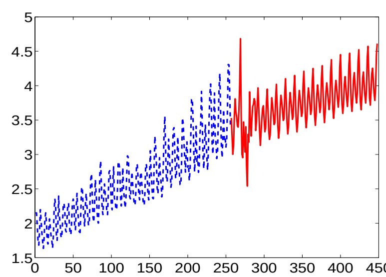
(Scenario 11)



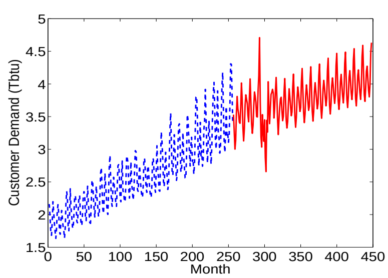
(Scenario 12)



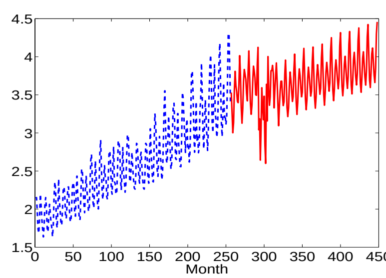
(Scenario 13)



(Scenario 14)

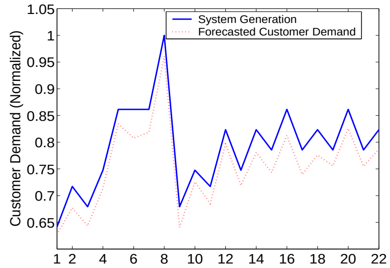


(Scenario 15)

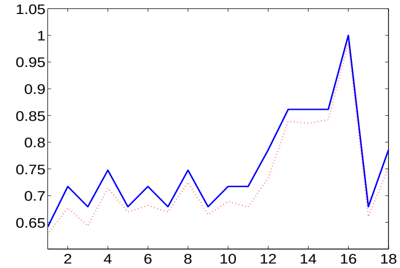


(Scenario 16)

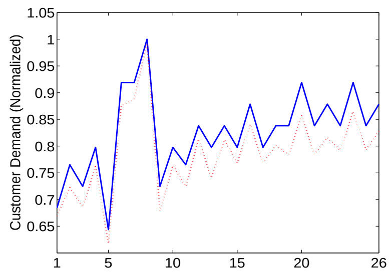
Figure 142: Customer Demand Forecasted Under Each Scenario (9-16)



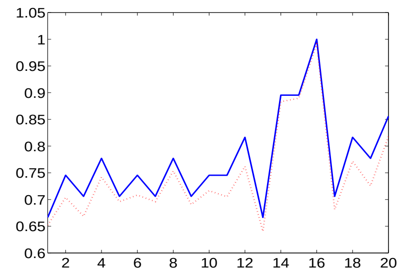
(Scenario 1)



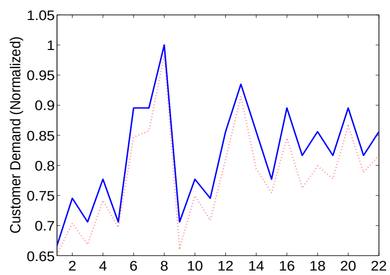
(Scenario 2)



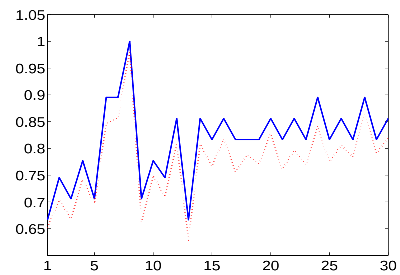
(Scenario 3)



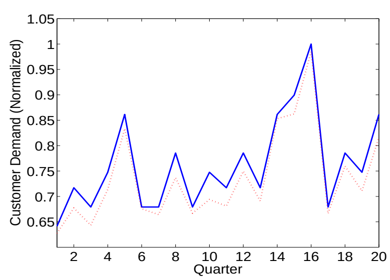
(Scenario 4)



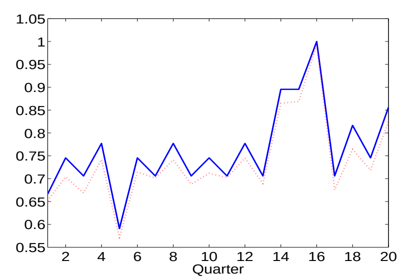
(Scenario 5)



(Scenario 6)

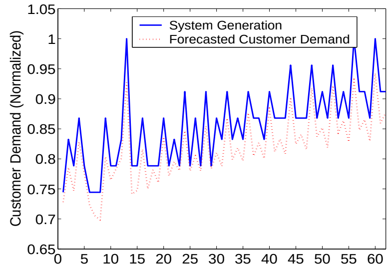


(Scenario 7)

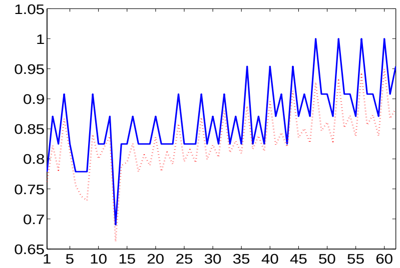


(Scenario 8)

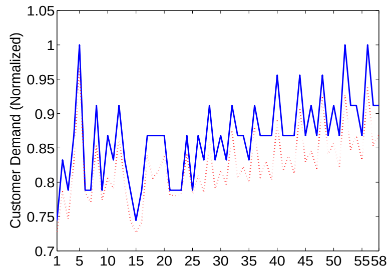
Figure 143: System Generation vs. Customer Demand Under Each Scenario (1-8)



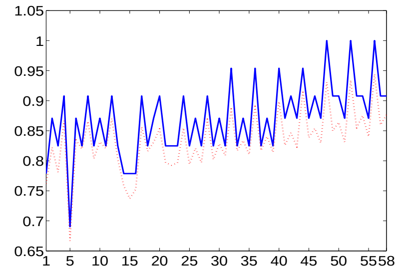
(Scenario 9)



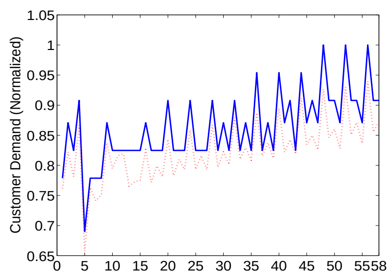
(Scenario 10)



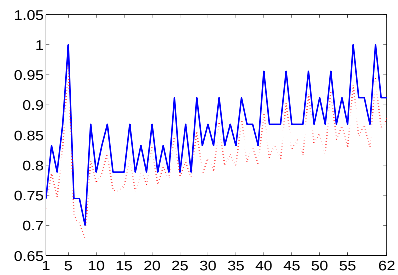
(Scenario 11)



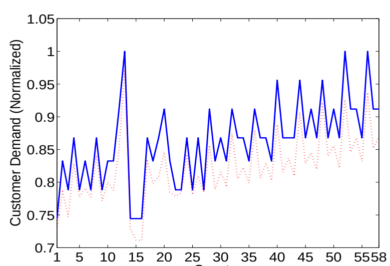
(Scenario 12)



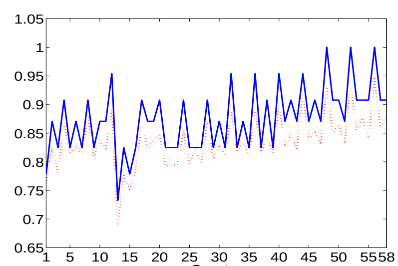
(Scenario 13)



(Scenario 14)

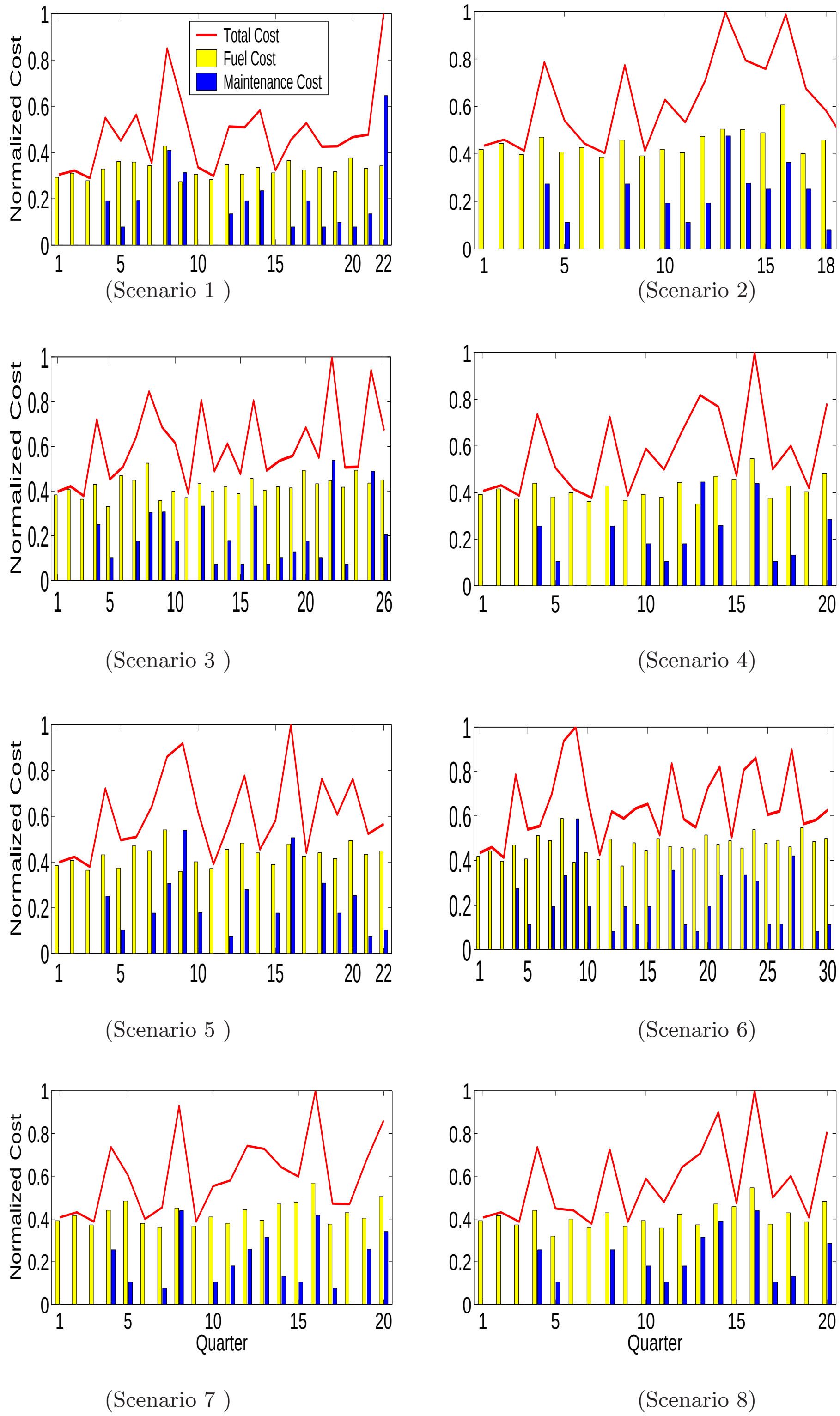


(Scenario 15)



(Scenario 16)

Figure 144: System Generation vs. Customer Demand Under Each Scenario (9-16)



(Scenario 1) (Scenario 2)

(Scenario 3) (Scenario 4)

(Scenario 5) (Scenario 6)

(Scenario 7) (Scenario 8)

Figure 145: System Total LCC Under Each Scenario (1-8)

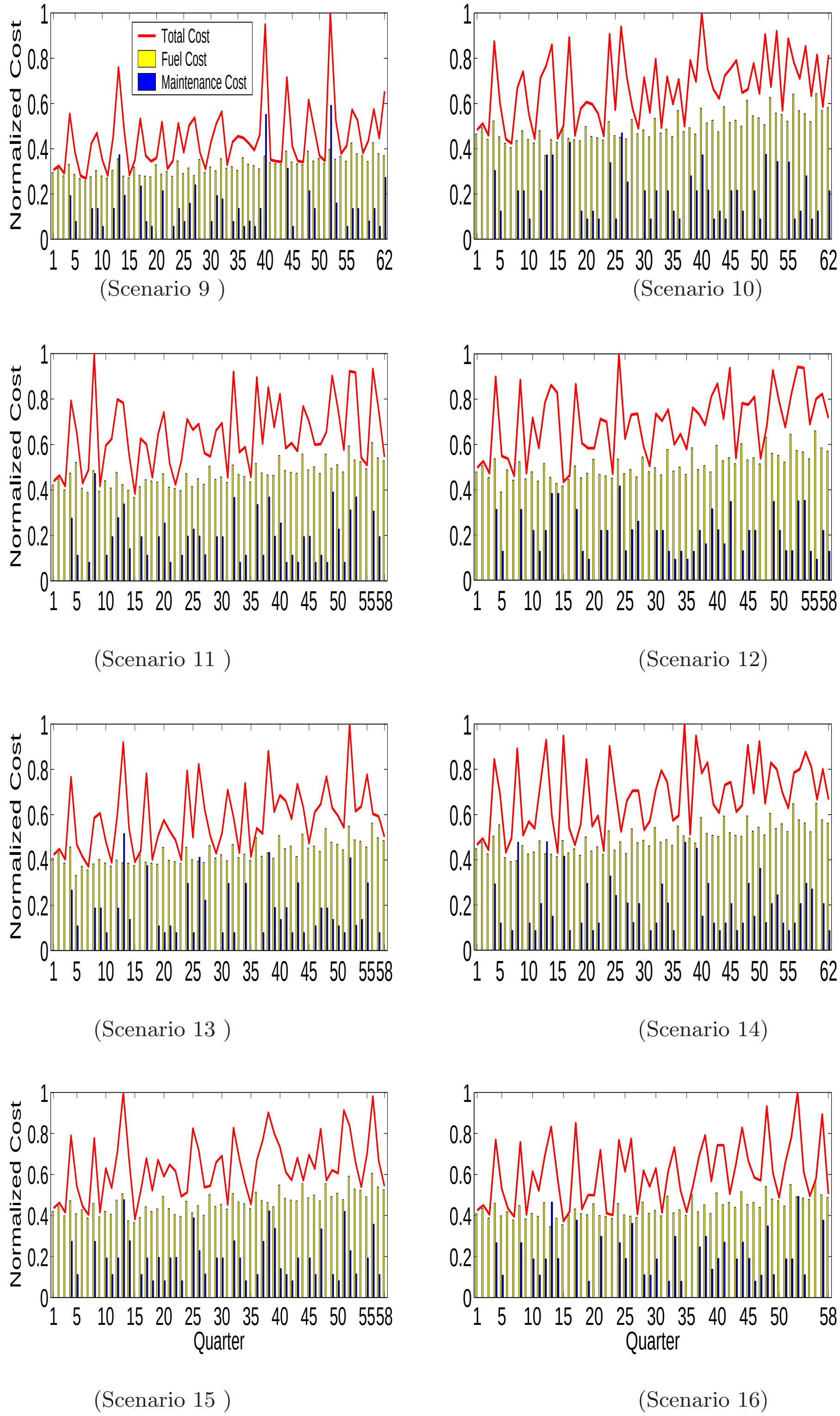


Figure 146: System Total LCC Under Each Scenario (9-16)

Table 53: Total Cost for Each Scenario

Scenario	1	2	3	4	5	6	7	8
Total Cost ($\mathcal{BLNV}$)	0.7781	0.5806	0.8788	0.6285	0.7487	1.0016	0.6596	0.6263
EOP (Q)	22	18	26	20	22	30	20	20
Cost/EOP ($\mathcal{BLNV}$)	0.03537	0.03226	0.03379	0.03143	0.03403	0.0339	0.03298	0.03132
Scenario	9	10	11	12	13	14	15	16
Total Cost ($\mathcal{BLNV}$)	2.0114	1.9182	1.8903	1.8146	1.7813	2.0216	1.8845	1.8286
EOP (Q)	62	62	58	58	58	62	58	58
Cost/EOP ($\mathcal{BLNV}$)	0.03244	0.03094	0.03259	0.03129	0.03071	0.03261	0.03249	0.03153

customer demand decreases very quickly. Hence, the EOP of the system extends to the far future. In this case, decision making should take place several times as the target is approaching and as more information is obtained. Such a rough estimate, nevertheless, is still useful in the preliminary stage of the uncertainty exploration.

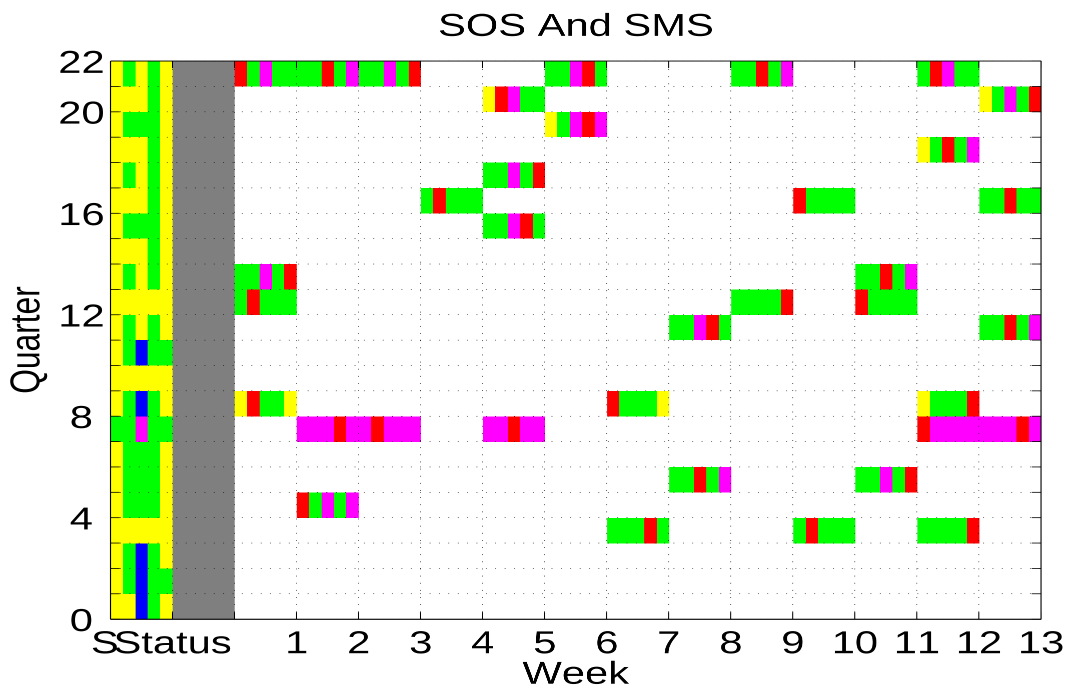


Figure 147: Scenario 1: SOS and SMS

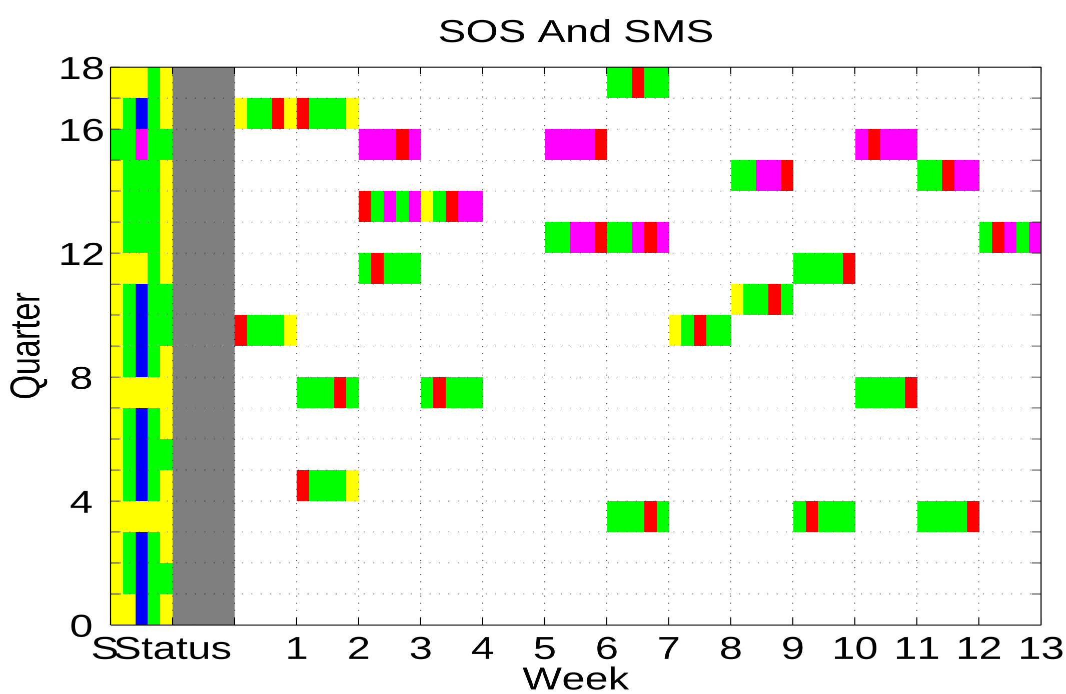


Figure 148: Scenario 2: SOS and SMS

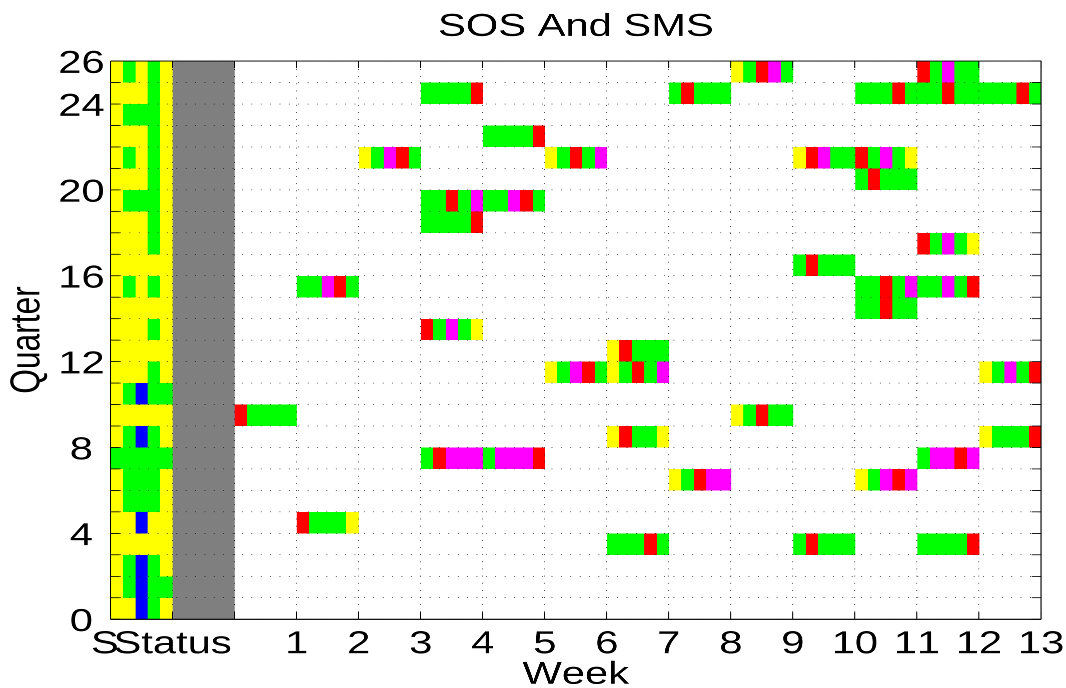


Figure 149: Scenario 3: SOS and SMS

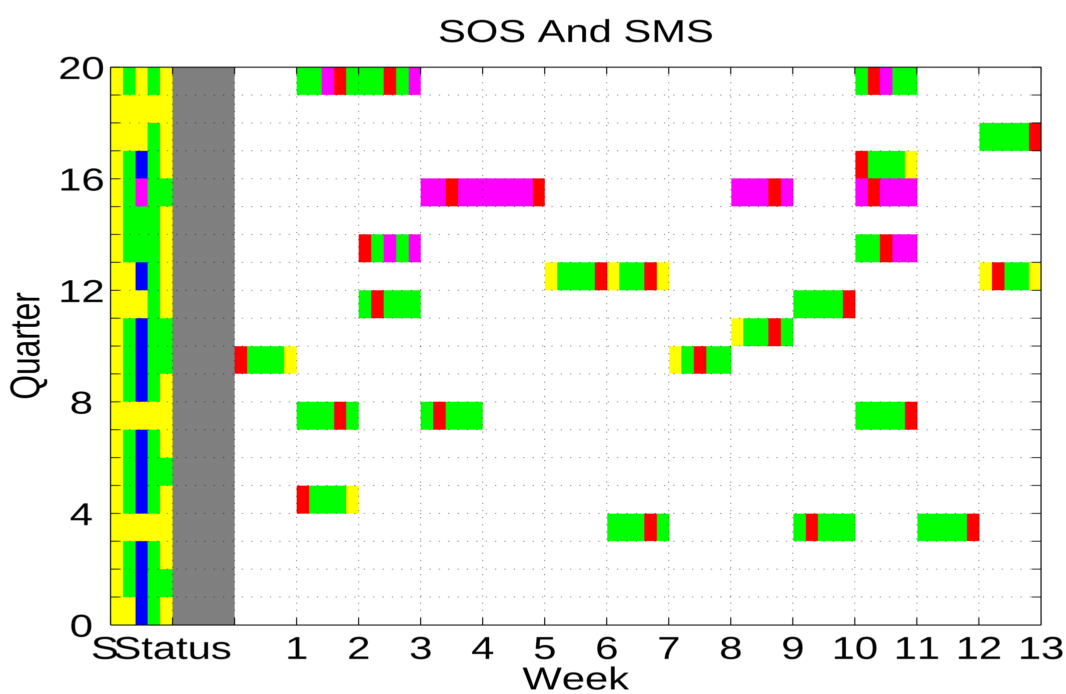


Figure 150: Scenario 4: SOS and SMS

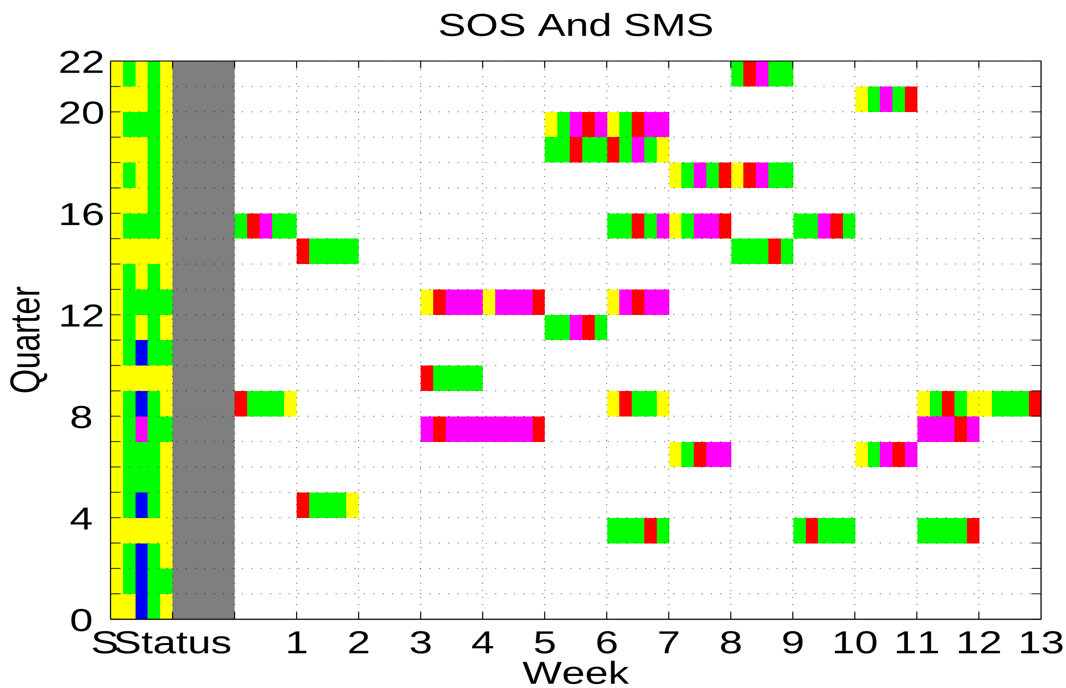


Figure 151: Scenario 5: SOS and SMS

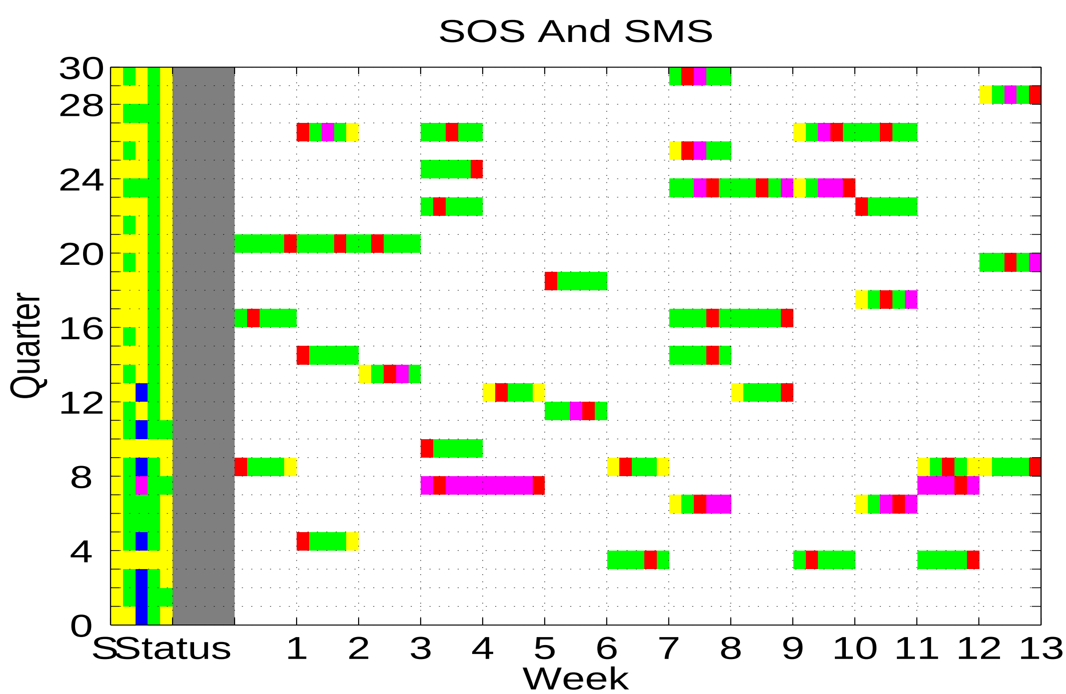


Figure 152: Scenario 6: SOS and SMS

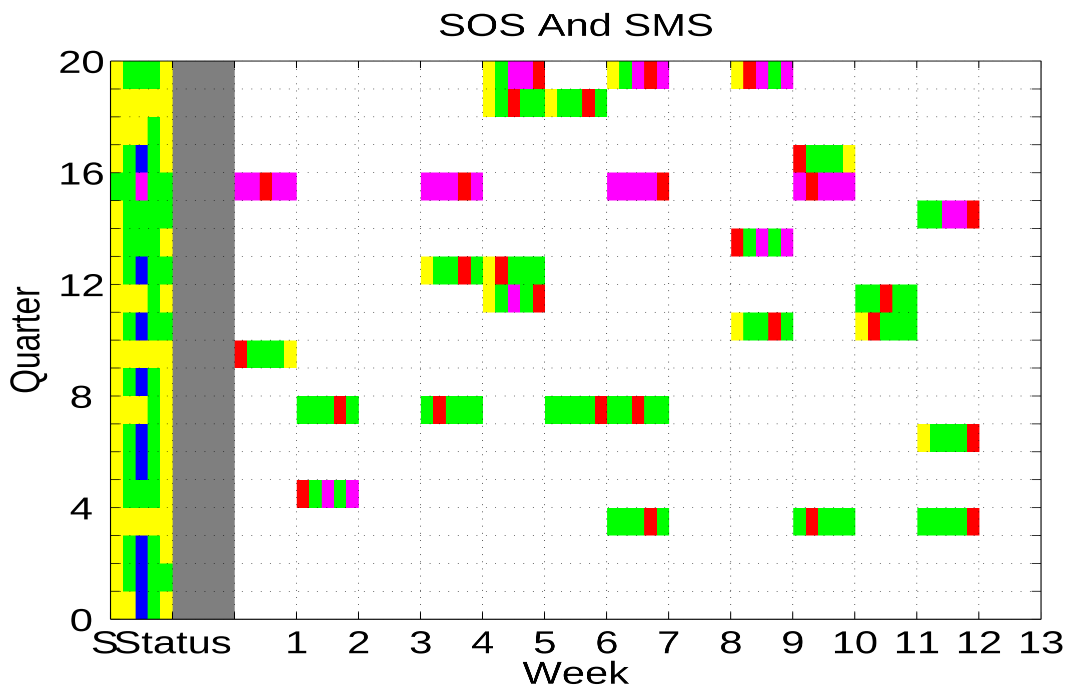


Figure 153: Scenario 7: SOS and SMS

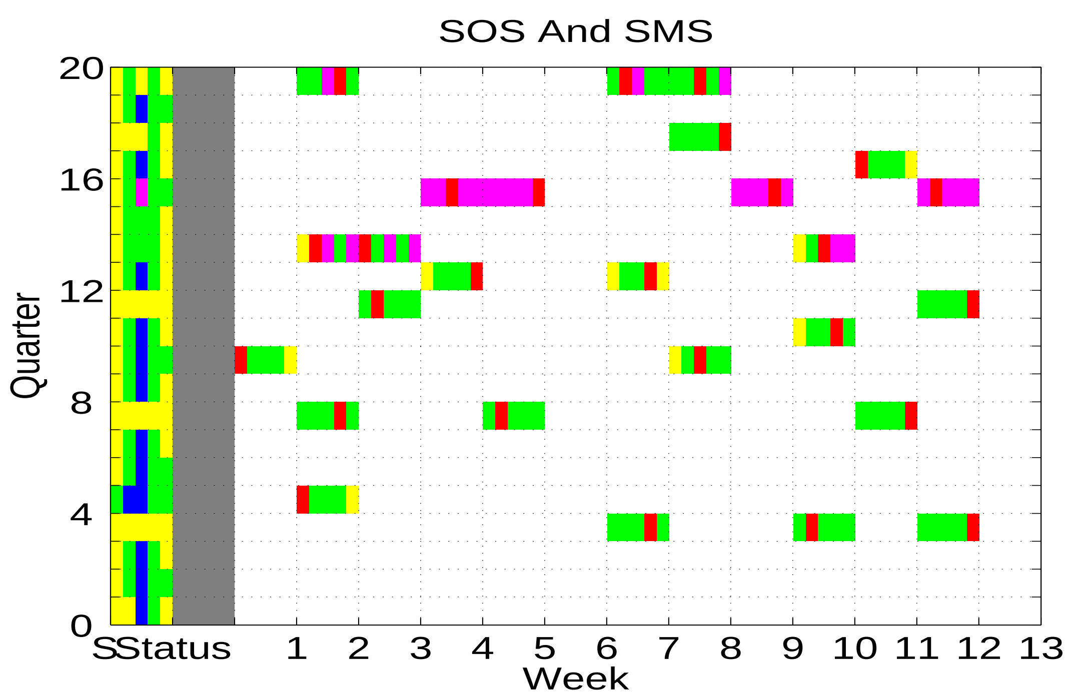


Figure 154: Scenario 8: SOS and SMS

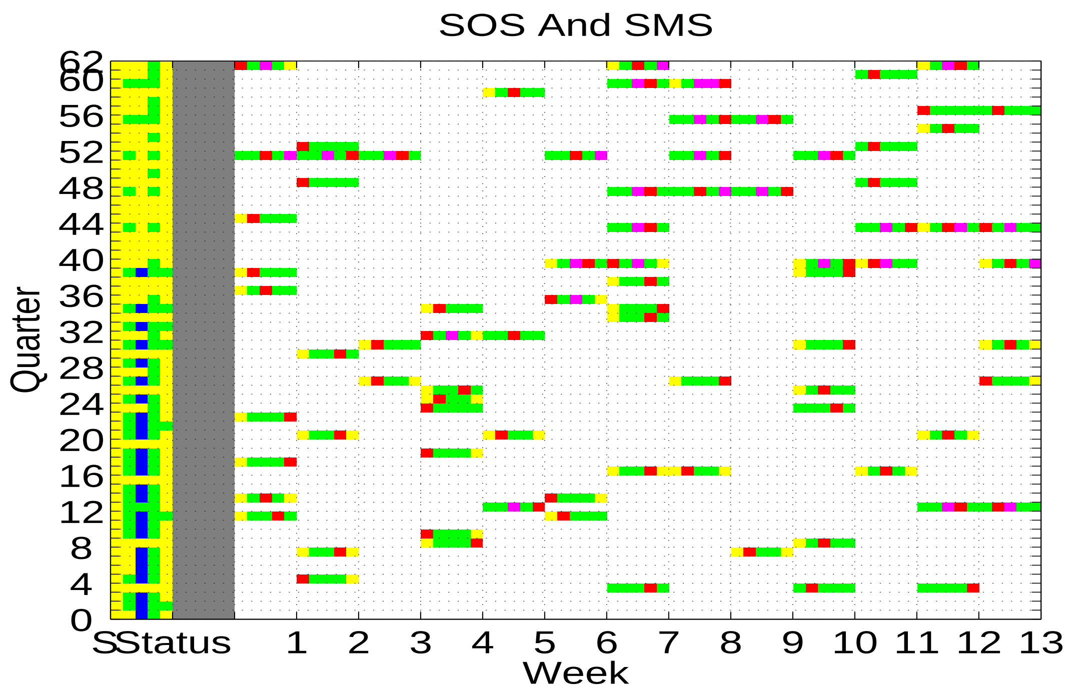


Figure 155: Scenario 9: SOS and SMS

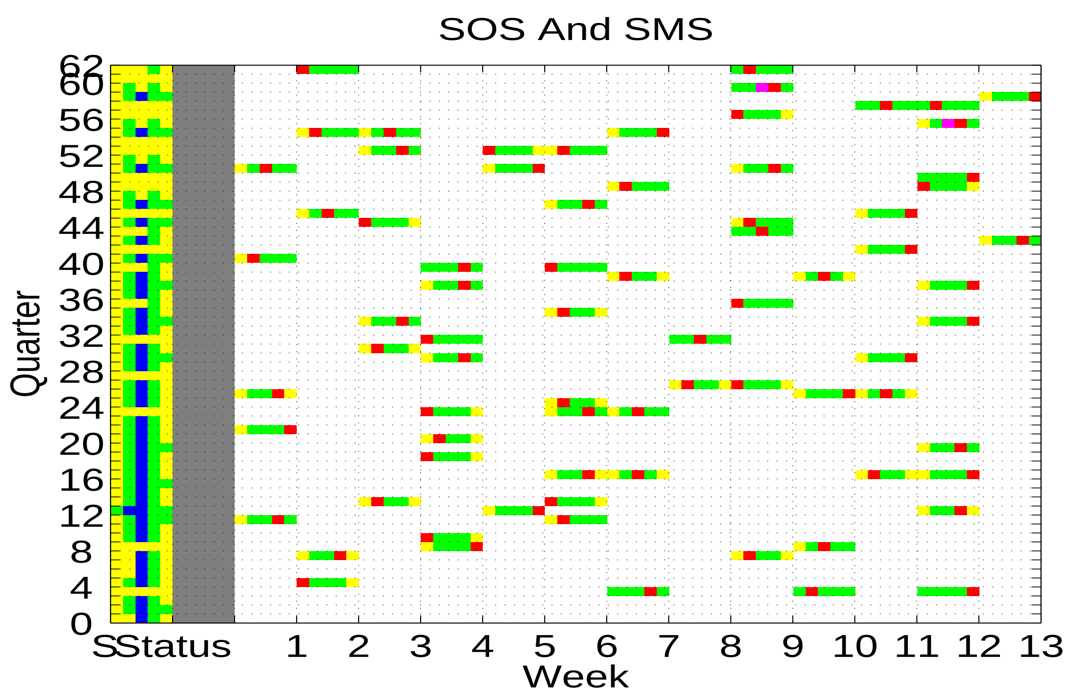


Figure 156: Scenario 10: SOS and SMS

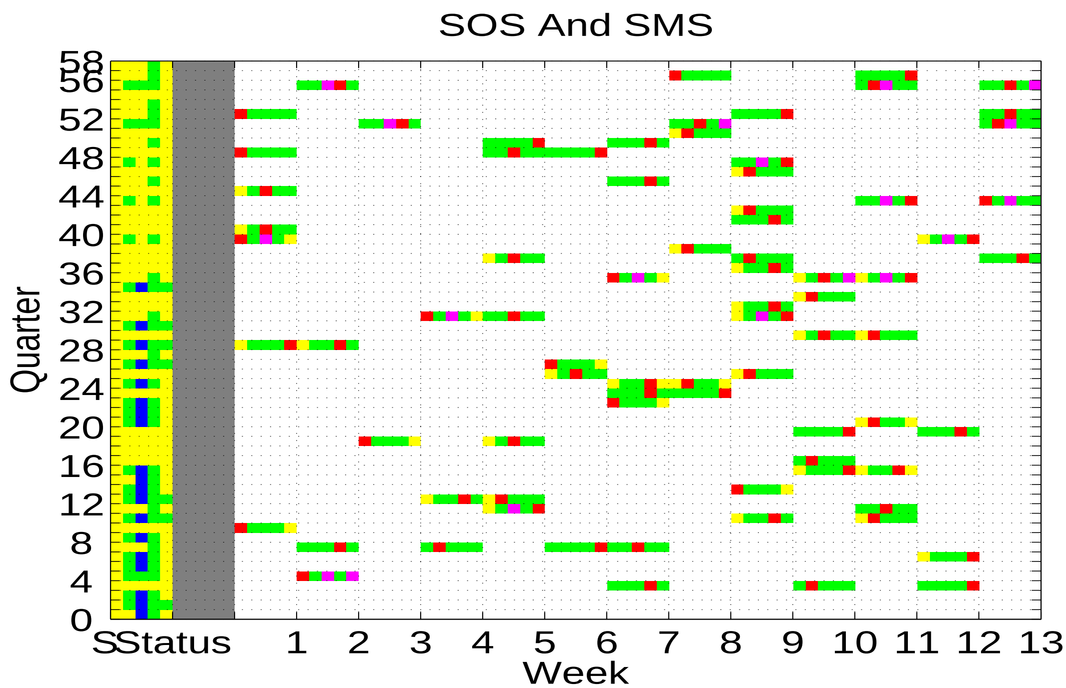


Figure 157: Scenario 11: SOS and SMS

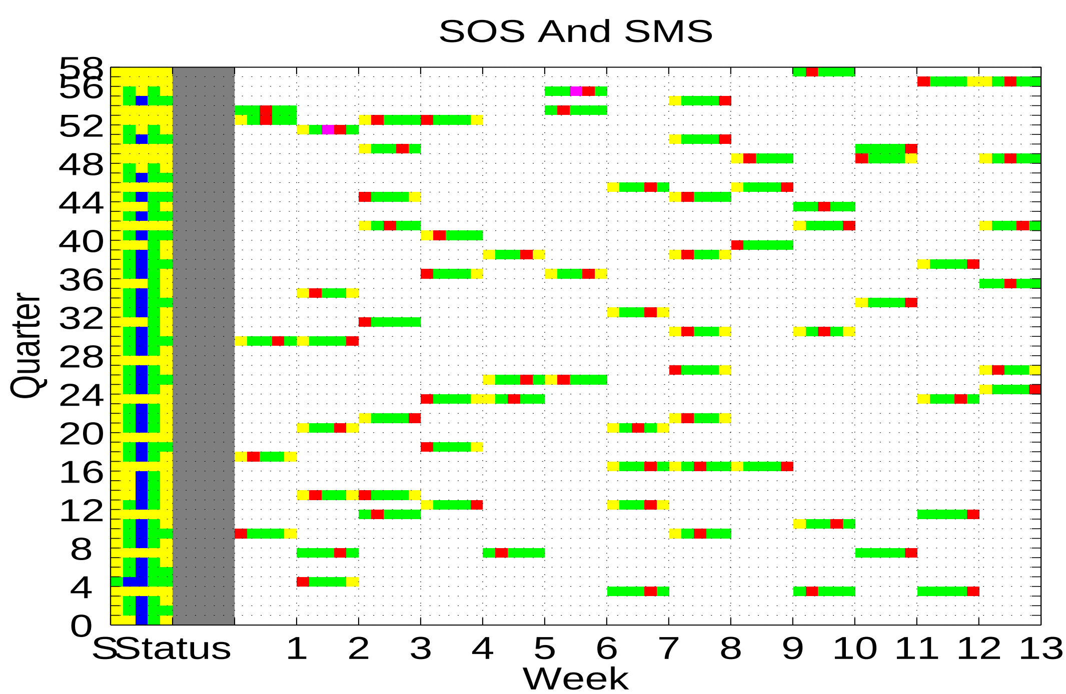


Figure 158: Scenario 12: SOS and SMS

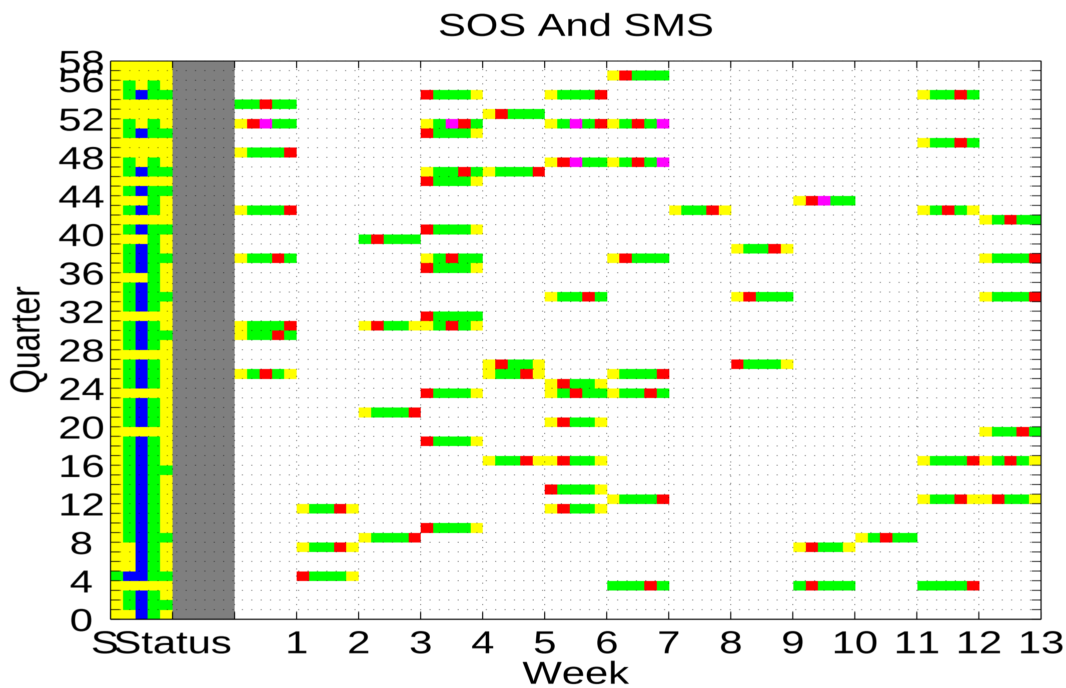


Figure 159: Scenario 13: SOS and SMS

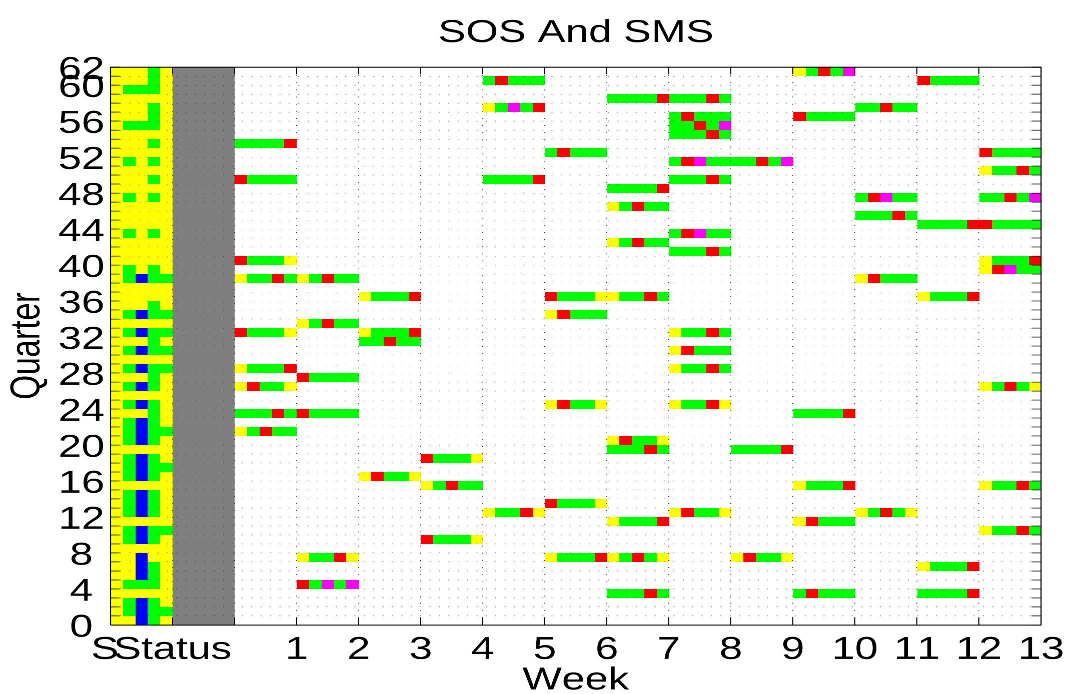


Figure 160: Scenario 14: SOS and SMS

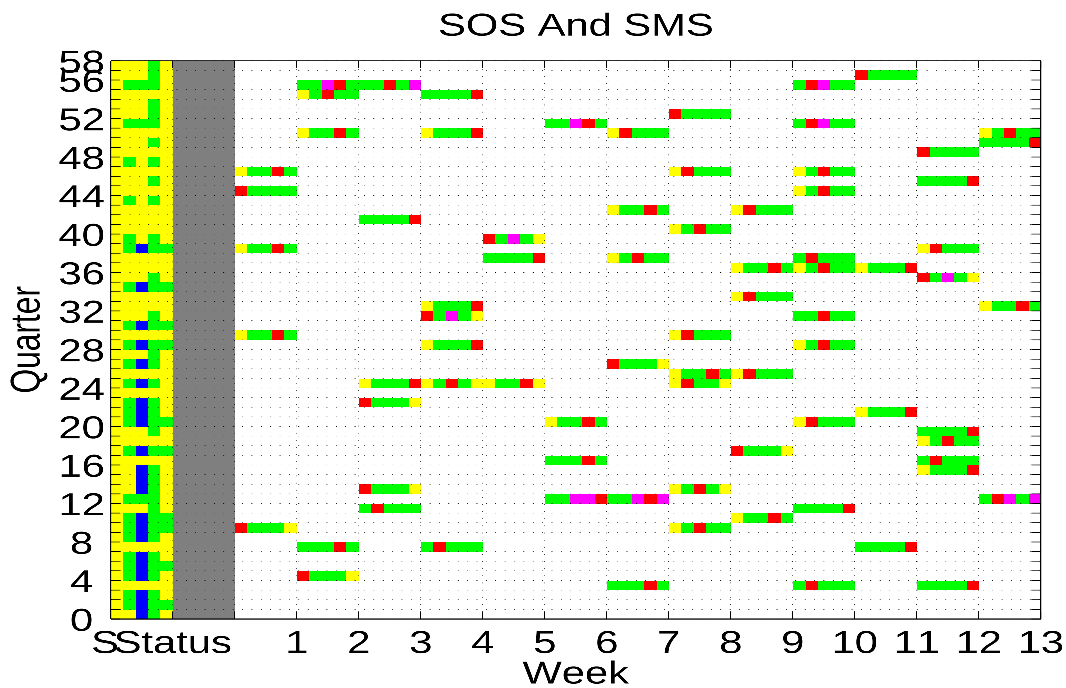


Figure 161: Scenario 15: SOS and SMS

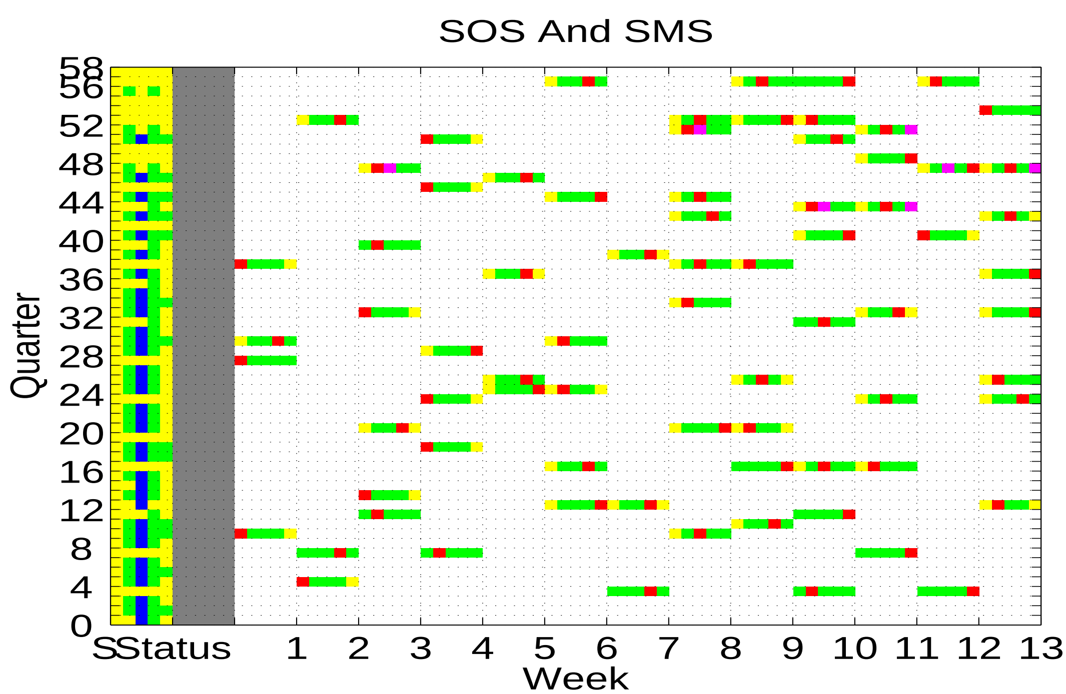


Figure 162: Scenario 16: SOS and SMS

CHAPTER VI

CONCLUSIONS

In this dissertation, a generic, system-level DM process has been developed, presented, and tested to manage diverse and often widely dispersed power generation units as a single, easily scaled and deployed fleet system in order to achieve true system excellence by fully utilizing the critical assets of a power producer. The development and presentation of this process is brought to a close in this chapter. Closure is sought by returning to the research questions posed in Chapter 1 and reviewing the answers that have been offered. Limitations of the research and possible avenues of future work are then discussed.

6.1 Conclusions

As stated in Chapter 1, the primary objective in this dissertation is to formulate a physics-based, system-level DM process that can help power plants reduce LCC and satisfy customer demand through improvements in both the forecasting methodology and the DM process. In particular, the improvements in the DM process required for multi-scale DM problems are exploited in the context of the following motivating research questions:

- *How will the cross-scale interactions be accounted for?*
- *How will the timescale for each decision action be determined?*
- *How will "point events" be handled?*

Answers and Thesis references to Research Questions:

- The major mid- to long-term decisions of an electric power plant include maintenance scheduling (SMS), operational planning (SOP) and capacity expansion planning (SCEP) on the system level. Figure 1 in Chapter 1 illustrates the time horizons for them and their interactions. When SOP is being considered, one problem that cannot be ignored is the SMS problem. These two decision actions are closely interrelated and discussed in Chapter 2. Maintenance activities aim at operating the system with a high level of reliability and security (see Section 2.1.3). However, the generation units under maintenance might contribute to lower system reserved capacity and higher production costs, leading to a tradeoff between how to appropriately commit and operate the generation units and how to schedule maintenance activities so that operating and maintenance cost can be minimized. Maintenance activities take place on the order of several hours to several weeks. Operation process has a time constant of years. Therefore, a dual timescale system that replaces the single time scale traditionally used in the power plant fleet management is utilized in Chapter 3. A large time scale is used for SOP and a fine time scale is needed for the description of SMS.

- Customer demand, electricity prices and natural gas prices, whose characteristics are illustrated in Figures 17, 18, and 21, respectively, are the main input to SOP. All these data series clearly have seasonal variations. The determination of SOS should optimally capture the seasonality in customer demand. A quarter of a year has been selected as the time step for SOP. Updating system status for each quarter is not only beneficial to capturing the seasonal characteristics of the historical data, but is also capable of operating power plants profitably without extra expenditure of generating too much power or not being able to satisfy customer demand. A week is selected as the time step for establishing the SMS, based on the fact that the maintenance window is usually in terms of weeks. Unit status is identified for each week. The lead time that a power plant needs to react to maintenance activities is at most a week. Therefore, it provides a systematic mechanism of dealing with maintenance activities. Implementation of this approach is demonstrated in Chapter 3 (see Section 3.1.4 in particular).

- System status is monitored for each week, see Section 3.1.6. Power plants update the system status for each quarter, see Section 3.1.5. "Point events," such as maintenance activities or special events, act as a trigger that switches to the use of the fine time scale, week. System status is adjusted to meet customer demand and minimize total cost in the maintenance window. Therefore, during the period that "point events" occur, power plants operate their generation units based on the time scale w in order to quickly update the system status and minimize the costs associated with the "point events." This approach helps zoom in on "point events" during long-term operation. It speeds up the response of power plants to the changes in both the power plant itself and the electric market. Therefore, it could facilitate the achievement of optimal operational conditions. The reactions of electric power plants to unscheduled maintenance activities are discussed in Section 5.2.2.

The development of the forecasting methodology is carried out in the context of the following motivating research questions:

- *How will data analysis be facilitated by utilizing MRA (e.g., NDWT)to extract critical information from historical data for forecasting?*
- *What available modeling techniques can be appropriately applied to each time scale? How will external information be incorporated into the forecasting process?*
- *How will the behavior of forecasting errors be identified?*

Answers and Thesis References to Research Questions: This process includes a forecasting system whereby the decision makers have the ability to make more informed decisions based on more accurate forecasting information through the use of MRA and the synergy of several modeling techniques properly combined at different time-scales.

- MRA analyzes data according to scale. It provides local representation of data in both the time and frequency domains. In MRA, a data series can be viewed as composed of a smooth background and details on top of it (see Section 3.2.2). This characteristic is utilized to de-trend and de-seasonalize a time series. Among the host of various wavelet transforms, the NDWT is identified as the most suitable one for tasks of forecasting in Section 3.2.3. The trend component is "located" in scaling coefficients and on coarse levels of detail (lower frequencies) as opposed to the high-frequency component, which requires fine-grained detail space for its description. The signature of the seasonal component is located at the intermediate levels. In this manner, by separating coarse, intermediate, and fine levels of detail, the time series may be de-trended, de-seasonalized, and de-noised in a mathematically logical way.

- For each level, a suitable technique for analyzing the data and making predictions is found. The main processes of modeling techniques are ARMAX model, harmonic regression, and Holt-Winters' method. The trend component is predicted using Holt-Winters' method. For the seasonal components, harmonic regression is used to make forecasts with estimated seasonal periods. The de-trended and de-seasonalized time series should have a stationary signature. Hence, the ARMA part of an ARMAX model should be able to describe this stationary high-frequency component, and, at the same time, the input of the ARMAX model will enable the model to take into account external inputs, which helps decision makers to easily assess and trade-off the impact of various external forces on the development and evolution of power plants. Thus, the high-frequency component filtered out by the wavelet technique can be fitted by an ARMAX model, which will be used to make forecasts for the high-frequency component in the sequel. The forecasting process, WAW, is illustrated in Section 3.3 and applied to forecast customer demand, natural gas prices, electricity prices in Chapter 4.

- The behavior of the forecasting errors during the wavelet transform and the inverse wavelet transform might have a significant impact on the accuracy of the forecasting. The behavior of forecasting errors has been investigated through comprehensive empirical analyses. Several scenarios have been chosen to explore the behavior of the WAW methodology in Section 3.3.2. Research has shown that forecasting errors are not magnified through the wavelet transform and inverse wavelet transform, and they can be estimated by auto-regressive (AR) processes in order to derive an additional systematic component to add back to the forecasting model. The order of AR processes are robust with respect to the type of wavelet filter used in the transform and log-linear with the length of the input data series.

The research questions that must be considered in the evaluation of the impact of the external business environment are as follows:

- *How will the bias of the estimate of the LCC needed to drive the business be evaluated?*
- *What are the critical sources of uncertainty and their features?*
- *How will the uncertainty from the external business environment be explored?*

Answers and Thesis References to Research Questions:

- An estimate of the LCC of driving the business for a power plant in baseline operation over the planning time horizon is provided. This is a point estimate of the LCC that the power plant will actually spend in the future. The block bootstrap method is used to measure the bias of the estimated system LCC to the actual LCC needed to drive business. Block bootstrap is a nonparametric method and was developed to approximate the sampling distribution and variance of statistics while preserving the internal structure of the data in Section 3.3.3. It is performed on the historical data to generate pseudo samples that are utilized as input to the forecasting method WAW. Each set of forecasting results based on the pseudo samples is used as an input to the DM process. An optimal operating strategy is chosen to achieve the minimal total LCC for each of them. A distribution of LCC based on these pseudo samples can be obtained. The bias is then calculated to be the difference between the baseline value and the average value of the distribution in Section 5.4.
- Two factors, weather and economic development, are identified as very important driving forces in the electric market in Section 3.4. Two indicators, W and E, which represent their functions have been chosen. Factor W is a phenomenon that occurs instantaneously and disappears instantaneously. Factor E occurs gradually and disappears instantaneously. For each factor, a vector is used to describe its condition: the first element representing the value of an phenomenon, the second representing

the time that this phenomenon occurs, and he third element representing the duration of the impact of this phenomenon. The ranges of each element of each factor should be identified, so one morphological field is required for each factor to determine its condition.

- Scenario analysis, acknowledging the uncertain business environment, considers a scenario to be a descriptive narrative of a set of relevant factors that describe alternative representations of future socio-economic conditions from a probabilistic point of view. Scenario analysis is utilized to identify the external factors, such as weather, economic development. Then scenarios are generated to describe the possible future conditions of electric power plants. Scenario analysis is carried out in Section 5.5 which provides possible future conditions of electric power plants.

A proof of concept investigation was performed on a typical power plant. The power plant was selected because it had been challenged by waves of change brought on by deregulation, globalization, and restructuring, and by the need for both critical assets that drive their business and accurate forecasting information on which to base the planning of the system activities and thus the performance of fleet management. In order to achieve system excellence, the power plant required more sophisticated fleet management approaches with more accurate forecasting support systems to manage diverse and often widely dispersed generating units as a single, easily scaled and deployed system. The proposed forecasting method WAW was first utilized to provide the forecasting information for customer demand, natural gas prices, and electricity prices via wavelet transforms, the ARMAX model and generic statistical methods. The forecasting results were validated with real data and compared with those of the traditional Holt-Winters' method. The results of the comparison showed that the forecasting method WAW proposed in this study can provide better overall performance and more accurate forecasting results.

Then the DM process was carried out by incorporating cross-scale interactions and forecasting information. First, the unit level conditions and then the system status were identified. An appropriate time scale for each decision action, such as SMS, SOP, and

SCEP, was identified, which helped the power plant to “zoom in” on “point events” so as to react quickly to any changes occurring in the system. The time scale for system operation planning was a quarter and for SMS, a week. The SOS, SMS, and SCEP were identified, and the distribution of LCC over the EOP was provided. The bias of the estimated total costs was calculated through the block bootstrap to measure on average the over/underestimates of the actual total cost.

These analyses were used as the baseline for an exploration of uncertainty. Scenario analysis was performed to construct a limited number of consistent and highly contrasting scenarios that might broadly cover the main possible evolutions of the system. The SOS, SMS, and total LCC distributions were identified for each scenario to prepare decision makers to face the uncertainties of the future as portrayed in the scenarios, and informed them of the potential impact of some key driving forces that might influence the future development of the power plant.

6.2 Future Work and Recommendations

The overall goal of this study was to approach the power plant fleet management problem from a system-level point of view based on accurate market information forecasting. The results of incorporating cross-scale interactions and applying the forecasting method WAW showed that this DM process is quite successful. However, some concerns regarding some components of this process should be addressed as this research continues to develop.

In the power plant fleet management, unit conditions were categorized based on the unit generation and the selection criterion, the ratio of unit generation to the FFH for generating that output. However, many other factors, such as environmental regulations that must be met by power plants, should be considered. A multi-property criterion should be developed for ranking the operating conditions. The inclusion of environmental constraints and other factors into the criterion for ranking the generation units in future power plant fleet management would be of special importance.

One major contributor to the total LCCs of power plants, with the exception of fuel costs, is maintenance costs. The calculation of maintenance costs is a comprehensive task that includes research from several fields. For example, the sum of the costs of inventory, ordering, shipping, and material can be minimized through a careful tradeoff between demand and inventory. A certain safety inventory should be carried out for the purpose of satisfying the demand for maintenance parts that exceeds the amount forecasted for a given period of time due to unscheduled maintenance or other special events. The determination of an appropriate level of safety inventory should consider two factors:

- The uncertainty of both demand and supply for maintenance resources
- The desired level of maintenance resources availability

Therefore, further insight into these two areas should provide a more accurate maintenance cost.

This research focuses on the study of the behavior of the electric market, and the response of power plants to it. The goal of the operation of power plants is to minimize the total cost while meeting customer demand. It is carried out by ignoring the interactions among multiple power plants and the resulting dynamics in the market environment. Actually, in any business, interactions with customers, suppliers, business partners, and competitors play an integral role in any decision and its consequences. Advances in information technology and e-commerce further enrich and broaden these interactions by increasing the degree of connectivity among the different parties involved in the commerce. Given that each system is part of a complex web of interactions, any business decision or action taken by a system affects the multiple entities that interact with or within that system, and vice versa. The strategic interaction of a system with its competitors, customers, and suppliers can be modeled as a game, and hence, game theory can be utilized to analyze it. To identify what decisions it must make, each system must understand how other systems or customers form their decisions and expectations. Given an understanding of the behavior of all the players, each one can then form its own best response decision. Therefore, how power plants operate within this complex web is a research area that might be worth putting effort into.

The block bootstrap was used in this research to measure the bias of the LCC of the system. With regard to the number of pseudo samples, no fixed answer to it was found, but an infinite number of replications, which the bootstrap requires on a formal level, might produce an accurate measurement. The key to the usefulness of the bootstrap is that it converges in terms of numbers of replications reasonably quickly, so running a finite number of replications should be sufficient, assuming that the number or replications were large enough. The above statement contains the key to choosing the right number of replications:

1. Choose a large but tolerable number of replications. Obtain the bootstrap estimates.

2. Change the random number seed. Obtain the bootstrap estimates again, using the same number of replications.

3. Determine whether the results reflect significant changes. If so, the first number you chose was too small, so try a larger number. If the results are similar, you probably have a large enough number. To be sure, perform the step 2 several more times.

The difficulty in performing block bootstrap in this research is the complexity of the forecasting method. The generated replications are used as the pseudo historical data to be input to the forecasting process. Wavelet transform is performed on each pseudo sample to partition it into different scale levels. Then for each level, a suitable technique is used to analyze the data and make predictions. ARMAX, harmonic regression, and Holt-Winters' method are used for the high-frequency component, the seasonal component, and the trend component, respectively. Different pseudo historical data require different number and values for model parameters in order to achieve the best overall forecasting results, so a relatively small number of pseudo samples was used in this study to roughly estimate the bias of the system total cost.

The original idea of the bootstrap was developed in [29] for approximating the sampling distribution and the variance of many statistics under the assumption of i.i.d data. To achieve this purpose, synthesis data are generated by independently re-sampling (with replacement) from the original observations, their statistics of interest are computed, and the variance among the replicas is used to estimate the sample variance. The extension to non i.i.d time series data is not trivial and it usually depends on both the structure of the time series (in [76] the case of stationary time series is considered) and the statistics of interest. To preserve the particular structure of the time series, block bootstraps, including the one used in this study, are often used. However the performance of these strategies depends on two competing constraints: faithfully reproducing the statistics of the original observations and producing sufficient variability among the surrogate series [4]. Recent efforts [71] for developing resampling methods for long memory processes typically transform the data into another domain (e.g., wavelets or a Fourier based domain) that maximizes the de-correlation among coefficients. Several wavelet surrogate methods have been proposed, see [14] and [17]. Therefore, in this study, transforming the time series of customer demand, natural gas prices, and electricity prices into the wavelet domain and then performing bootstrap might be another way to estimate the bias of the total life cycle costs.

In the scenario analysis, weather and economic development were identified as the two main factors that contributed most significantly to the forecasting process and consequently to the DM process. In this study, two different types of external factors that act similarly in some way to weather and economic development were utilized. The impact of weather is usually short-term, and the impact of economic development is more gradual and long-term. In order to evaluate the impact of these external factors, a more accurate forecasting of these factors in the future is a must. Unfortunately, significantly accurate forecasting for weather or economic development is not available. The approach used in this research, however, is sufficient for a preliminary uncertainty exploration and for the purpose of bounding uncertainty for power plants. Further research along this line could prove very useful.

APPENDIX A

THE COMPUTATIONS OF MAINTENANCE FACTORS

$$\text{Maintenance Interval (hr)} = \frac{24000}{\text{Maintenance Factor}}$$

Where:

$$\text{Maintenance Factor} = \frac{\text{Factored Hours}}{\text{Actual Hours}}$$

Factored Hours = (K + M * I) * (G + 1.5 D +A H + 6 P)

Actual Hours = (G + D + H +P)

- **G = Annual Based Load Operating Hours on Gas Fuel**
- **D = Annual Based Load Operating Hours on Distillate Fuel**
- **H = Annual Operating Hours on Heavy Fuel**
- **A = Heavy Fuel Severity Factor (Residual A = 3 to 4, Crude A = 2 to 3)**
- **P = Annual Peak Load Operating Hours**
- **I = Percent Water/Steam Injection Referenced to Inlet Airflow**
- **M & K = Water/ Steam Injection Constants**

M	K	Control	Steam Injection	N2/N3 Material
0	1	Dry	< 2.2%	GTC-222FSX-414
0	1	Dry	> 2.2%	GTD-222
0.18	0.6	Dry	> 2.2%	FSX-414
0.18	1	Wet	> 0%	GTD-222
0.55	1	Wet	> 0%	FSX-414

Figure 163: Hot-Gas-Path Inspection: Hours-Based Criterion

$$\text{Maintenance Interval (st)} = \frac{S}{\text{Maintenance Factor}}$$

Where:

$$\text{Maintenance Factor} = \frac{\text{Factored Starts}}{\text{Actual Starts}}$$

Factored Starts = 0.5 NA + NB + 1.3 NP +20 E + 2 F + $\Sigma_i (a_i T_i)$

Actual Starts = NA + NB +NP + E + F + T

- **S = Max. Starts Based Maintenance Interval (Model Size Dependent)**
- **NA = Annual Number of Part Load Start/Stop Cycles (< 60% Load)**
- **NB = Annual Number of Normal Base Load Start/Stop Cycles**
- **NP = Annual Number of Peak Load Start/Stop Cycles**
- **E = Annual Number of Emergency Starts**
- **F = Annual Number of Fast Load Starts**
- **T = Annual Number of Trips**
- **a = Trip Severity Factor = f(Load)**
- **n = Number of Trips Categories (i. e. Full Load, Part Load, etc.)**

Model Series	MS6B/MS7EA	MS6FA	MS9E	MS7F/7FA/9F/9FA
S	1200	900	900	900

Figure 164: Hot-Gas-Path Inspection: Starts-Based Criterion

$$\textbf{Maintenance Interval (hr)} = \frac{144000}{\textbf{Maintenance Factor}}$$

Where:

$$\textbf{Maintenance Factor} = \frac{\textbf{Factored Hours}}{\textbf{Actual Hours}}$$

Factored Hours = H + 2 P + 2 TG
Actual Hours = H + P

H = Base Load Hour

P = Peak Load Hours

TG = Hours on Turning Gear

Figure 165: Rotor Inspection: Hours-Based Criterion

$$\textbf{Maintenance Interval (st)} = \frac{5000}{\textbf{Maintenance Factor}}$$

Where:

$$\textbf{Maintenance Factor} = \frac{\textbf{Factored Starts}}{\textbf{Actual Starts}}$$

Factored Starts = Fh * Nh + Fw1 * Nw1 + Fw2 * Nw2 + Fc * Nc + Ft * Nt
Actual Starts = Nh + Nw1 + Nw2 + Nc + Nt

	Fast	Normal
Fh = Hot Start Factor (Down 1-4 Hrs)	1.0	0.5
Fw1 = Warm 1 Start Factor (Down 4-20 Hrs)	1.8	0.9
Fw2 = Warm 2 Start Factor (Down 20-40 Hrs)	2.8	1.4
Fc = Cold Start Factor (Down > 40 Hrs)	4.0	2.0
Ft = Trip From Load Factor	4.0	4.0
	PG 7241 PG 9351 Designs	

Nh = Number of Hot Starts

Nw1 = Number of Warm 1 Starts

Nw2 = Number of Warm 2 Starts

Nc = Number of Cold Starts

Nt = Number of Trips

Figure 166: Rotor Inspection: Starts-Based Criterion

$$\text{Maintenance Factor} = \frac{\text{Factored Hours}}{\text{Actual Hours}}$$

Factored Hours = Σ_i (Ki * Afi * Api * ti), 1 to n Operating Modes

Actual Hours = Σ_i (ti)

i = Discrete Operating Modes or Operating Practice or Time Interval

ti = Operating Hours at Load in a Given Operating Mode

Api = Load Severity Factor

Ap = 1.0 Up to Base Load

Ap = exp(0.018 * Peak Firing Temp Adder in Deg F) for Peak Load

Afi = Fuel Severity Factor (Dry)

Af = 1.0 for Gas Fuel

Af = 1.5 for Non-DLN (or 2.5 for DLN) for Distillate Fuel

Af = 2.5 for Crude (Non-DLN)

Af = 3.5 for Residual (Non-DLN)

Ki = Water/Steam Injection Severity Factor

(% Steam Referenced to Inlet Airflow, w/f = water to fuel ratio)

K = Max(1.0, exp(0.34(%Steam-2.00%))) for Steam, Dry Control Curve

K = Max(1.0, exp(0.34(%Steam-1.00%))) for Steam, Wet Control Curve

K = Max(1.0, exp(1.8(w/f-0.8))) for Water, Dry Control Curve

K = Max(1.0, exp(1.8(w/f-0.4))) for Water, Wet Control Curve

Figure 167: Combustor Inspection: Hours-Based Criterion

$$\text{Maintenance Factor} = \frac{\text{Factored Starts}}{\text{Actual Starts}}$$

Factored Starts = Sum(Ki * Afi * Ati * Api * Asi * Ni), i=1 to n Operating Modes

Actual Starts = Sum(Ni)

- **i = Discrete Start/Stop (or Operating Practice)**
- **Ni = Start/Stop Cycles in Given Operating Mode**
- **Asi = Start Type Severity Factor**
  - **As = 1.0 for Normal Start**
  - **As = 1.2 for Start with Fast Load**
  - **As = 3.0 for Emergency Start**
- **Ati = Trip Severity Factor = 0.5 + (exp(0.0125*%Load)) for Trip**
- **Api = Load Severity Factor**
  - **Ap = 1.0 Up to Base Load**
  - **Ap = exp(0.09*Peak Firing Temp Adder in Deg F) for Peak Load**
- **Afi = Fuel Severity Factor (Dry)**
  - **Af = 1.0 for Gas Fuel**
  - **Af = 1.25 for Non-DLN (or 1.5 for DLN) for Distillate Fuel**
  - **Af = 2.5 for Crude (Non-DLN)**
  - **Af = 3.5 for Residual (Non-DLN)**
- **Ki = Water/Steam Injection Severity Factor**
  - **(% Steam Referenced to Inlet Airflow, w/f = water to fuel ratio)**
  - **K = Max(1.0, exp(0.34(%steam-2.00%))) for Steam, Dry Control Curve**
  - **K = Max(1.0, exp(0.34(%Steam-1.00%))) for Steam, Wet Control Curve**
  - **K = Max(1.0, exp(1.8(w/f-0.8))) for Water, Dry Control Curve**
  - **K = Max(1.0, exp(1.8(w/f-0.4))) for Water, Wet Control Curve**

Figure 168: Combustor Inspection: Starts-Based Criterion

REFERENCES

[1] Akansu, A. N. and Haddad, R. A., *Multiresolution Signal Decomposition: Transforms, Subbands, and Wavelets.* Academic Press.

[2] Alawaji, S. H. and Lo, K. L., "A new approach for solving the problem of unit commitment," in *2nd International Conference on Advances in Power System Control, Operation and Management, APSCOM-93*, pp. 571 – 577, Dec. 1993.

[3] Amjady, N., "Short term hourly load forecasting using time series modeling with peak load estimation capability," *IEEE Transactions on Power Systems*, vol. 16, pp. 798 – 805, Nov. 2001.

[4] Angelini, C., Cava, D., Katul, G., and Vidakovic, B., "Resampling hierarchical processes in the wavelet domain: A case study using atmospheric turbulence," *Elsevier Science*, 2004. Submitted.

[5] Archer, G. and Giovannoni, J. M., "Statistical analysis with bootstrap diagnostics of atmospheric pollutants predicted in teh apsis experiment," *Water, Air, and Soil Pollution*, vol. 106, pp. 43 – 81, 1996.

[6] Arsham, H., "Time-critical decision making for economics and finance." http://home.ubalt.edu/ntsbarsh/.

[7] Australian, "National electricity market." http://www.nemmco.com.au. Management Company Limited (NEMMCO).

[8] Bailer, A. J., Oris, J. T., Lange, N., Ryan, L., Billard, L., Brillinger, D., Conquest, L., and Greehouse, J., "Assessing toxicity of pollutants in aquatic systems," *Biometry*, pp. 25 – 40, 1994.

[9] Beran, R. and Ducharme, G. R., *Asymptotic Theory for Bootstrap Methods in Statistics.* Universitee de Montreal, Canada: Centre de Recherches Mathematiques, Nov. 1991.

[10] Bhavaraju, M. P., Hebson, J. D., and Wood, W., "Emerging issues in power system planning," in *Proc. of the IEEE*, vol. 40, pp. 891–898, June 1989.

[11] Billinton, R. and Allan, R. N., *Reliability Assessment of Large ELectric Power System.* Kluwer International Series in Engineering & Computer Science, Kluwer Academci/Plenum Publishers, 1998.

[12] Billinton, R., Li, Y. W., and Li, W., *Reliability Assessment of ELectrical Power System Using Monte Carlo Methods.* Kluwer Academci/Plenum Publishers, 1994.

[13] Borison, A. and Mueller, H., "Forecasting fuel requirements uncertainty," *IEEE Transaction on Engineering Management*, vol. 3, pp. 1046 – 1051, Aug. 1988.

[14] Breakspear, M., Brammer, M., and Robinson, P. A., "Construction of multivariate surrogate sets from nonlinear data using the wavelet transform," *Physica D*, vol. 182, no. 1 – 2, pp. 1 – 22, 2003.

[15] Breipohl, A. M., "Electricity price forecasting models," in *Power Engineering Society Winter Meeting, IEEE*, vol. 2, pp. 963 – 966, Jan. 2002.

[16] Brockwell, P. J. and Davis, D. V., *Introduction to Time Series and Forecasting.* Springer-Verlag, 2002.

[17] Bullmore, E., Long, C., Suckling, J., Fadili, J., Calvert, G., Zelaya, F., Carpenter, T. A., and Brammer, M., "Colored noise and computational inference in neurophysiological time series analysis: Resampling methods in time and wavelet domains," *Hum Brain Sapp*, vol. 12, no. 2, pp. 61 – 78, 2001.

[18] Carlstein, E., "The use of subseries values for estimating the variance of a general statistic from a stationary sequence," *Ann. Statist.*, vol. 14, pp. 1171 – 1179, 1986.

[19] Chen, H. M., Vidakovic, B., and Mavris, N. D., "Multiscale forecasting method using armax models," *Technological Forecasting and Social Change*, 2004. Accepted.

[20] Cho, M. Y., Hwang, J. C., and Chen, C. S., "Customer short term load forecasting by using arima transfer function method," in *International Conference on Energy Managemetn and Power Delivery, Proc. of EMPO '95*, vol. 1, pp. 317 – 322, Nov. 1995.

[21] Chopra, S. and Meindl, P., *Supply Chain Management: Strategy, Planning and Operation.* Prentice-Hall, second ed., 2004.

[22] Chowdhury, B. H. and Rahman, S., "A review of recent advances in economic dispatch," *IEEE Transactions on Power Systems*, vol. 5, pp. 1248 – 1259, Nov. 1990.

[23] Christianse, W. R., "Short term load forecasting using general exponential smoothing," *IEEE Transactions on Power Systems*, pp. 900 – 910, 1971.

[24] Chui, C. K., *Wavelets: A Tutorial in Theory and Applications.* Wavelet Analysis and its Applications, Volume 2, Academic Press, 1992.

[25] Cohen, A. and Ryan, R. D., *Wavelets and Mutiscale Signal Processing.* Applied Mathematics & Mathematical Computation, Chapman & Hall, 1995.

[26] Cooley, R. L., "Confidence intervals for groundwater models using linearization, likelihood, and bootstrp methods," *Ground Water*, vol. 35, pp. 869 – 880, 1997.

[27] Crousillat, E., "Risk and uncertainty in power planning." UNDP Gnereal Review Seminar, dec 1988.

[28] Davison, A. C. and Hinkley, D. V., *Bootstrap Methods and Their Application.* Cambridge University Press, 1997.

[29] Efron, B., "Bootstrap methods: Another look at the jackknife," *Annual Statistics*, vol. 7, pp. 1 – 26, 1979.

[30] Efron, B. and Tibshirani, R., *An Introduction to the Bootstrap*. New York: Chapman & Hall, 1993.

[31] Energy, R., "A dhirubhai ambani enterprise/glossary." http://www.rel.co.in/KnowledgeCenter/.

[32] Eydeland, A. and Wolniec, K., *Energy and Power Risk Managment: New Developments in Modeling, Pricing, and Hedging*. John Wiley & Sons, Inc., 2003.

[33] Filho, M. B. D. C., Silva, A. M. L., Arieti, V. L., and Ribeiro, S. M. P., "Probabilistic load modeling for power system expansion planning," in *3rd International Conference on Probabilistic Methods Applied to Electric Power Systems*, pp. 203 – 217, July 1991.

[34] Flardrin, P., *Time-Frequency/Time Scale Analysis*. Wavelet Analysis and its Applications, Academic Press, 1999. Translated from French by J. Stockler.

[35] Fosso, O. B., Gjelsvik, A., Haugstad, A., Mo, B., and Wangensteen, I., "Generation scheduling in a deregulated system: The norwegian case," *IEEE Transactions on Power Systems*, vol. 14, pp. 75 –81, Feb. 1999.

[36] Frauendorfer, K., Glavitsch, H., and Bacher, R., *Optimization in Planning and Operation of Electric Power Systems*. Physica-Verlag, A Springer-Verlag Company, 1992.

[37] Fukuyama, Y. and Chiang, H. D., "A parallel genetic algorithm for generation expansion planning," *IEEE Transactions on Power Systems*, vol. 11, pp. 955 –961, May 1996.

[38] Good, P. I., *Resampling Methods: A Practical Guide to Data Analysis*. Birkhauser, 1999.

[39] Gorenstin, B. G., Campodonico, N. M., Costa, J. P., and Pereira, M. V. F., "Power system expansion planning under uncertainty," *IEEE Transactions on Power Systems*, vol. 8, pp. 129 –136, Feb. 1993.

[40] Granger, C. W. J. and Newbold, P., *Forecasting Economic Time Series*. New York: Academic Press, 1977.

[41] Graps, A., "An introduction to wavelets." http://www.amara.com/IEEEwave/IEEEwavelet.html.

[42] Hall, P., *The Bootstrap and Edgeworth Expansion*. New York: Springer-Verlag, 1992.

[43] Hardle, W., Kerkyacharian, G., Picard, D., and Tsybakov, A., *Wavelets Approximation and Statistical Applications*. Springer, 1998.

[44] Hoeft, R., Janawitz, J., and Keck, R., "Heavy-duty gas turbine operating and maintenance considerations," tech. rep., GE Energy Services, Atlanta, GA, 2001.

[45] Irizarry, R. A., "Local regression with meaningful parameters," *The American Statistician*, vol. 55, no. 1, pp. 72 – 79, 2001.

[46] Jia, N. X., Yokoyama, R., and Zhou, Y. C., "A novel approach to long term load forecasting where functional relations and impact relations coexist," in *International Conference on Electric Power Engineering, PowerTech Budapest 99*, pp. 38 –, Aug. 1999.

[47] Kahn, H. and Wiener, A. J., *The Year 2000.* New York: Macmillan, 1967.

[48] Kaiser, G., *A Friendly Guide to Wavelets.* Library of Congress Cataloging-in-Publication Data, Lowell, MA: Birkhauser, 1994.

[49] Kehlhofer, R., Bachmann, R., Nielsen, H., and Warner, J., *Combined Cycle Gas & Steam Turbine Power Plants.* Library of Congress Cataloging-in-Publication Data, Tulsa, Oklahoma: Pennwell Publishing Company, second ed., 1999.

[50] Kim, Y. C. and Ahu, B. H., "Multicriteria generation-expansion planning with global environmental considerations," *IEEE Transaction on Engineering Management*, vol. 40, pp. 154 – 161, May 1993.

[51] Kirby, R. M. and Mavris, N. D., "Forecasting technology uncertainty in preliminary aircrat design," in *1999 World Aviation Conference*, (San Francisco), Oct. 1999.

[52] Koether, A. B., Snyder, R. D., and Ord, J. K., "Forecasting models and prediction interval for the multiplicative holt-winters' method," *International Journal of Forecasting*, vol. 2, no. 0, pp. 269 – 286, 2001.

[53] Kunsch, H., "The jackknife and the bootstrap for general stationary observations," *Annals of Statistics*, vol. 17, pp. 1217–1241, 1989.

[54] Lahiri, S. N., *Resampling Methods for Depedent Data.* Springer Series in Statistics, Springer, 2003.

[55] Leou, R. C., "A flexible unit maintenance scheduling considering uncertainties," *IEEE Transactions on Power Systems*, vol. 16, pp. 552 –559, Aug. 2001.

[56] Liang, R. H. and Cheng, C. C., "Combined regression-fuzzy approach for short-term load forecasting," in *IEE Proc. Generation, Transmission and Distribution*, vol. 147, pp. 261 – 266, July 2000.

[57] Linares, P., "Multiple criteria decision making and risk analysis as risk management tools for power systems planning," *IEEE Transactions on Power Systems*, vol. 17, pp. 895 – 900, 2002.

[58] Liu, R. Y. and Singh, K., *Moving Blocks Jackknife and Bootstrap Capture Weak Dependence*, pp. 225–248. Exploring the Limits of Bootstrap, New York: Wiley, 1992.

[59] Lowis, A. K., Maab, P., and Rieder, A., *Wavelets: Theory and Applications.* John Wiley & Sons, 1997.

[60] Martino, B. D., Fusco, G., Mariani, E., Randino, R., and Ricci, P., "A medium and short term load forecasting model for electric industry," in *Power Industry Computer Applications Conference, PICA-79, IEEE Conference Proceedings*, pp. 186 – 191, May 1979.

[61] Modis, T., *Predictions: Society's Telltale Signature Reveals the Past and Forecasts the Future*. New York: Simon&Schuster, 1992.

[62] Moon, Y. H., Lee, J. D., and Lee, T. S., "A new economic dispatch algorithm for thermal unit generation scheduling in power system," in *Power Engineering Society Winnter Meeting, IEEE*, vol. 2, pp. 1034 – 1039, Jan. 2000.

[63] Mukerji, R., Merrill, H. M., Erickson, B. W., Parker, J. H., and Friedman, R. E., "Power plant maintenance scheduling: Optimizing economics and reliability," *IEEE Transactions on Power Systems*, vol. 6, pp. 473 – 483, May 1991.

[64] Muller, K. R., Kohlmorgen, J., and Plawelzik, K., "Analysis of switching dyanmics with computing neural networks," tech. rep., Univ. Tokyo, Tokyo, Japan, 1997.

[65] Nogales, F. J., Contreras, J., Conejo, A. J., and Espinola, R., "Forecasting next-day electricity prices by time series models," *IEEE Transactions on Power Systems*, vol. 17, pp. 342 – 348, May 2002.

[66] Operator, I. E. M., "10 year outlook, ontario demand forecast, from january 2004 to december 2013," tech. rep., Independent Electricity Market Operator, 2003. IMO_REP_0098v2.0.

[67] Oppel, L. J. and Aronson, D., "Information technology, its increased importance in the power industry after deregulation," in *Power Engineering Society Summer Meeting, IEEE*, pp. 215 – 221, July 1999.

[68] Papalexopoulos, A. D. and Hesterberg, T. C., "A regression-based approach to short-term system load forecasting," *IEEE Transactions on Power Systems*, vol. 5, pp. 1535 – 1547, Nov. 1990.

[69] Parlos, A. G., Oufi, E., Muthusami, J., Patton, A. D., and Atiya, A. F., "Development of an intelligent long-term electric load forecasting system," in *Proc. ISAP '96. International Conference on Intelligent Systems Applications to Power Systems*, vol. 2, pp. 288–292, Jan. 1996.

[70] Peng, H., Ozaki, T., Toyoda, Y., and Oda, K., "Modeling and control of systems with signal dependent nonlinear dyanmics," in *Proc. the European Control Conference*, pp. 42 – 47, 2001.

[71] Percival, D. B., Sardy, S., and Davison, A. C., *Wavestrapping Time Series: Adaptive Wavelet-Based Bootstrapping*. Cambridge University Press, 2001. Edited by W. J. Fitzgerald, R. L. Smith, A. T. Walden and P. C. Young.

[72] Percival, D. B. and Walden, A. T., *Wavelet Methods for Time Series Analysis*. Cambridge University Press, 2000.

[73] Polikar, R., "The engineer's utimate guide to wavelet analysis." http://users.rowan.edu/ polikar/WAVELETS.

[74] Politis, D. and Romano, J. P., *A Circular Block Resampling Procedure for Stationry Data*, pp. 263–270. Exploring the Limits of Bootstrap, New York: Wiley, 1992.

[75] Politis, D. N. and Romano, J. P., *A Nonparametric Resampling Procedure for Multivariate Confidence Regions in Time Series Analysis*, pp. 98–103. Computing Science and Statistics,Proceedings of the 22nd Symposium on the Interface, (Connie Page and Raoul LePage, eds.), Springer Verlag, 1992.

[76] Politis, D. N. and Romano, J. P., “The stationary bootstrap,” *Journal Amer. Statist. Assoc.*, vol. 89, no. 428, pp. 1303 – 1313, 1994.

[77] Praca, I., Ramos, C., and Vale, Z. A., “Competitive electricity markets: Simulation to improve decision making,” in *Power Tech Proc. 2001 IEEE Porto*, pp. 6 –, Sept. 2001.

[78] Prina, J., “Investment decision making in a deregulated electric industry using stochastic domnance,” in *Engineering Management Society, Proc. of the 2000 IEEE*, pp. 546–551, Aug. 2000.

[79] R. J. Hyndman, A. B. K., Snyder, R. D., and Grose, S., “A state space framework for automatic forecasting using exponential smoothing methods,” *International Journal of Forecasting*, vol. 18, no. 3, pp. 439 – 454, 2000.

[80] Raffaele, Z., Gambelli, D., and Vairo, D., “he future of organic farming in europe: A scenario analysis,” in *Proceedings of the International Symposium “Organic Agriculture in the Mediterranean basin*, 2001.

[81] Rajan, C. C. A. and Mohan, M. R., “An evolutionary programming-based tabu search method for solving the unit commitment problem,” *IEEE Transactions on Power Systems*, vol. 19, pp. 277 – 585, Feb. 2004.

[82] Rizk, N. J., “A genetic algorithm for economic dispatching of generators,” in *Proc. 2004 International Conference on Information and Communication Technologies: From Theory to Applications*, pp. 533 – 543, Apr. 2004.

[83] Schlosser, W. and Ebel, E., “Evaluating the time value of historical testing information in animal populations.” http://www.riskworld.com/Profsoci/ps5me002.htm, june 2004. Society for Risk Analysis.

[84] Senjyu, T., Takara, H., Uezato, K., and Funabashi, T., “One-hour-ahead load forecasting using neural network,” *IEEE Transactions on Power Systems*, vol. 17, pp. 113 – 118, Feb. 2002.

[85] Shahidehpour, M., “Investing in expansion: The many issues that cloud transmission planning,” *Power and Enegery Magazine, IEEE*, vol. 2, pp. 14 –18, Jan. 2004.

[86] Sheble, G. B. and Fahd, G. N., “Unit commitment literature synopsis,” *IEEE Transactions on Power Systems*, vol. 9, pp. 128 – 135, Feb. 1994.

[87] Silberglitt, R. and Hove, A., “Scenario analysis,” in *E-Vision 2000 Conference*, 2000. Supplementary Materials: Papers and Analyses.

[88] Silva, A. M. L. L., Anders, G. J., and Manso, L. A. F., “Generator maintenance scheduling to maximize reliability and revenue,” in *Power Tech Proc. 2001 IEEE*, pp. 6 –, Sept. 2001.

[89] Simar, L. and Wilson, P. W., "Sensitivity analysis of efficienty scores: How to bootstrap in nonparametric frontier models," *Management Science*, vol. 44, pp. 49 – 61, 1995.

[90] Skantze, P., Ilic, M., and Chapman, J., "Stochastic modeling of electric power prices in a multi-market environment," in *Power Engineering Society Winter Meeting, IEEE*, vol. 2, pp. 1109 – 1114, Jan. 2000.

[91] Sohtaoglu, N. H., "The effects of economic parameters on power transmission planning," in *9th Mediterranean Electrotechnical Conference*, vol. 2, pp. 941–945, May 1998.

[92] Sparling, D., "Forecasting." http://www.uoguelph.ca/ dsparlin/forecast.htm.

[93] Stedinger, J. R. and Taylor, M. R., "Synthetic streamflow generation 1. model verification and validation," *Water Resources Research*, vol. 18, no. 4, pp. 909 – 918, 1982.

[94] Stedinger, J. R. and Taylor, M. R., "Synthetic streamflow generation 2. effect of parameter uncertainty," *Water Resources Research*, vol. 18, no. 4, pp. 919 – 924, 1982.

[95] Stoft, S., *Power System Economics: Designing Markets for Electricity*. Wiley Interscience, 2002.

[96] Stoll, G. H., "Creating owner's competitive advantage through contractual services," tech. rep., GE Power Systems, Schenectady NY, 2001.

[97] Sullivan, T. G., "How to prepare a financial forecast." http://www.flexstudy.com. American Management Association.

[98] Tabari, N. M., Rahjbar, A. M., and Sadati, N., "Promoting the optimal maintenance schedule of generating facilities in open systems," in *International Conference on Power System Technology, Proc. PowerCon 2002*, pp. 641 –645, Oct. 2002.

[99] Tarnurr, N. R. and Stanton, L. W., *Quantitative Forecasting Methods*. The Duxbury Series in Statistics and Decision Sciences, Boston: PWS-KENT Publishing Company, 1989.

[100] Taylor, J. W. and Buizza, R., "Neural network load forecasting with weather ensemble predictions," *IEEE Transactions on Power Systems*, vol. 17, pp. 626 – 632, Aug. 2002.

[101] Teive, R. C. G., Silva, E. L., and Fonseca, L. G. S., "A cooperative expert system for transmission expansion planning of electrical power systems," *IEEE Transactions on Power Systems*, vol. 13, pp. 636 – 642, May 1998.

[102] Valenzuela, J., "A seeded memetic algorithm for large unit commitment problmes," *Journal of Heuristics*, vol. 8, pp. 173 – 195, Mar. 2002.

[103] Valsan, S. P. and Swarup, K. S., "Hopfiel neural network approach to the solution of economic dispatch and unit commitment," in *Proc. International Conference on Intelligent Sensing and Information Processing*, pp. 311 – 316, 2004.

[104] Vidakovic, B., *Statistical Modeling by Wavelets.* John Wiley & Sons, Inc., 1999.

[105] Vogel, R. M. and Stedinger, J. R., "The value of stochastic streamflow models in over-year reservoir design applications," *Water Resources Research*, vol. 24, no. 9, pp. 1483 – 1490, 1998.

[106] Walonick, D. S., "An overview of forecasting methodology," tech. rep., Professional Survey Software Web and Paper Questionnaires, 2004. http://www.statpac.com/.

[107] Wang, P. Y. and Wang, G. S., "Power system load forecasting with ann and fuzzy logic control," in *Proc. TENCON '93 IEEE Region 10 Conference on Computer, Commuication, Control and Power Engineering*, pp. 359 – 362, Oct. 1993.

[108] Wang, Y. and Handschin, E., "Unit maintenance scheduling in open systems using genetic algorithm," in *Transmission and Distribution Conference, 1999 IEEE*, pp. 334 –339, Apr. 1999.

[109] Wehrens, R. and Linden, W. E. V., "Bootstrapping principal component regression models," *Journal of Chemometrics*, vol. 11, pp. 157 – 171, 1997.

[110] Wheelwright, C. S. and Maridakis, S., *Forecasting Methods for Management.* Wiley Series on Systems and Controls for Financial Management, Wiley Publishers, 1973.

[111] Wu, F. F. and Varaiya, P., "Coordinated multilateral trades for electric power networks: Theory and implementation 1," *Electrical Power and Energy Systems*, vol. 21, June 1995.

[112] Xiao, J., Zhang, Y., and Wang, C. S., "A study and implementation of an intelligent load forecasting support system," in *Proc. PowerCon 2002 International Conference on Power System Technology*, vol. 2, Oct. 2002.

[113] Yellen, J., Al-khamis, T. M., Vemuri, S., and Lemonidis, L., "A decomposition approach to unit maintenance scheduling," *IEEE Transactions on Power Systems*, vol. 7, pp. 726 – 733, May 1992.

[114] Zhang, Q. and Benveniste, A., "Wavelet networks," *IEEE Transactions on Neural Networks*, vol. 3, pp. 889 – 898, Nov. 1992.

[115] Zhao, M., Kermanshahi, B., Yasuda, K., and Yokoyama, R., "Fuzziness of decision making and planning parameters for optimal generation mix," in *Electrotechnical Conference, MELECON '96., 8th Mediterranean*, pp. 215 – 221, May 1996.

[116] Zhao, Y. J., Chen, H. M., Waters, M., and Mavris, N. D., "Modeling and cost optimization of combined cycle heat recovery generator systems," in *2003 ASME Turbo Expo*, (Atlanta, GA), June 2003.

VITA

Hongmei Chen was born in Heilongjiang Province, China, on July 16, 1975. She graduated from Beijing University of Aeronautics & Astronautics, China, with a Bachelor of Science (B.Sc) degree in Aerospace Engineering in July 1998. Then in the same year, she was recommended for graduate study in the Department of Jet Propulsion of Beijing University of Aeronautics & Astronautics, China, and earned a Master of Science (M.Sc) degree in April 2001. In August 2001, she came to Georgia Tech and joined the Ph.D. program in the Aerospace System Design Laboratory in the School of Aerospace Engineering. She earned a second Master of Science degree in December 2002. Her primary research focus includes the development of methods that forecast market information and facilitate understanding of its impact on system behavior and properties and the system-level strategic decision-making process; the Bayesian approach to dealing with uncertainty and risk in decision analysis; and the operation and maintenance optimization of power plants